A GENOMIC AND GENETIC ANALYSIS OF DOUBLESEX TARGETS AND FUNCTION IN *DROSOPHILA* SEXUAL DIMORPHISM

By

Erin Alisa Jimenez

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
April, 2015

ABSTRACT

Sex determination pathways are diverse throughout the animal kingdom, but converge upon conserved genes that encode products that regulate sexual dimorphism. One such downstream factor across many diverged sex determination pathways is the *Drosophila doublesex* (*dsx*) gene. The role of *doublesex* is highly conserved in different insects and *dsx* homologs (*dsx*, *mab-3* related transcription factors, DMRTs) play roles in sexual differentiation in a diverse array of metazoans. In *Drosophila*, nearly all manifestations of sexual dimorphism between males and females are regulated by *doublesex*, yet there are only three known direct targets of DSX, which cannot account for the differences in regulation by DSX in sexually dimorphic tissues.

To gain a comprehensive understanding of DSX targets, we undertook multiple experimental approaches that allowed us to identify genes that were bound by DSX, genes whose expression changed in response to DSX perturbation, and genes that function in *dsx*-expressing cells. DSX protein binding was assayed by ChIP-seq and DamID-seq on S2 cells expressing tagged DSX^M or DSX^F. We also examined DSX occupancy in adult fat body and gonads using DamID-seq or DamID-chip. These experiments identified 3,717 genes bound by DSX in at least one occupancy dataset. Strikingly, we found that genes with the highest levels of DSX occupancy were bound by DSX in all occupancy data sets. This suggests one main mechanism of DSX action would be binding to potential targets in all tissues/contexts rather than having context-dependent targets. In this model of DSX action, additional inputs (such as segmental identity) would be needed to enact transcriptional regulation of bound genes in the appropriate context. Further strengthening this model, although 2,668 genes are bound by DSX in our adult

fat body occupancy data, less than 1% of these occupied genes show large and robust transcriptional changes in response to acute changes in DSX isoform.

We found that predicted DSX targets are significantly enriched in genes that yield phenotypes in sexually dimorphic tissues after RNAi knockdown in *dsx*-expressing cells (p=0.002). 41 (70.7%) of high probability DSX targets had phenotypes in at least one sexual dimorphic tissue compared to 7 (31.8%) of low probability targets. Altogether, the occupancy, transcriptional profiling, and functional testing have provided a detailed description of how dsx regulates sexual development.

New *dsx*-interacting genes include genes involved in insect hormone signaling. We have identified the *Ecdysone receptor* gene as a target of DSX. Since the *Drosophila* gonad represents an excellent model to dissect how DSX acts on a particular time and place to promote development of a sexually dimorphic tissue, we examined the *Ecdysone receptor* gene, which is involved in ecdysteroid signaling, for roles in gonad sexual development. My data supports the hypothesis that the steroid hormone ecdysone elicits a different response in the male vs. female gonad and that this difference is regulated by dsx and may be important for proper formation of the ovary vs. the testis. Rather than being strictly a genetic process, results from our experiments may demonstrate that sexual differentiation in the gonad occurs through a combination of signals that include sex specific hormone signaling. Since the formation of the gonad may represent processes that are conserved from flies to man, this research will provide insight into conserved genes that regulate developmentally similar pathways whose outcome generates major differences observed between the sexes.

Advisor: Dr. Mark Van Doren

Second Reader: Dr. Xin Chen

ACKNOWLEDGMENTS

I acknowledge the many people who made this thesis work possible. I foremost would like to thank my thesis advisor Dr. Mark Van Doren, whose guidance has been invaluable to this research and my growth as a scientist over the last five years. I also want to acknowledge past and present members of the Van Doren laboratory, especially Dr. Cale Whitworth who has been a mentor and important collaborator on the DSX occupancy and expression analyses, and also Kelly Baxter for assistance with the ecdysone reporter construct generation. I acknowledge Shekerah Primus, my friend and lab member, who has always been willing to discuss aspects of my research and who kindly read this thesis. Much of this work could not have been accomplished without the collaborative work of the DSX consortium. I am indebted to Dr. Brian Oliver, Dr. Stephen Goodwin, and Dr. Teresa Przytycka as well as scientists from their lab that contributed to this work: Dr. David Sturgill, Dr. Ryan Dale, Dr. Harold Smith, Dr. Megan Neville, Dr. Yoo-Ah Kim, Dr. Hania Pavlou, Dr. Zhen-Xia Chen, and Dr. Leonie Hempel. I especially thank Dr. Emily Clough who lead the DSX consortium and was always willing to teach, answer questions, and discuss aspects of the project. I acknowledge the faculty, staff, and other members of the Johns Hopkins community who aided in my growth as a scientist, especially my thesis committee members Dr. Xin Chen, Dr. Allan Spradling, Dr. Daniella Drummond-Barbosa, and Dr. Nick Ingolia. Michael McCaffery and Erin Pryce in the Johns Hopkins Integrated Imaging Center played an integral part in this work.

Only a few fly stocks have been generated from this work. The majority of the fly stocks and reagents were contributed from researchers around the world. The Bloomington Drosophila Stock Center and the Developmental Studies Hybridoma Bank were essential in contributing fly

stocks and antibodies.

Finally, I acknowledge my family and friends, all of whom supported and believed in me during my time at Johns Hopkins. I especially thank my mother, whom I dedicate this dissertation to for encouraging me, believing in my quest to be a scientist, and for making this achievement possible.

DEDICATION

This thesis is dedicated to my mother, Laurie.

For all your hard work and sacrifice you made so I can explore my strengths,

define my dreams and pursue them.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

In nature we find incredible differences in anatomical structures, physical appearance, and behavior between males and females. Sex-specific characteristics, or sexual dimorphism, are those traits that distinguish one sex from another, morphologically and behaviorally. But of all the sex-specific characteristics we may find in nature, perhaps the most outstanding sexual dimorphism observed is that of the gonad. Given that it is required in order to support spermatogenesis and oogenesis, sexual dimorphism of the gonad is arguably the most important. Therefore it is of great interest and importance to study the genetic and developmental mechanisms involved in proper formation of the gonad in the appropriate sex.

Mechanisms of sex determination

Diverse sex-determination mechanisms are utilized by species throughout the animal kingdom to create sexual dimorphism in the gonad. In mammals, insects, and birds sex is determined on a genetic basis. In other species such as reptiles, sex is determined by environmental cues, and in a variety of fish, sex is also based on social interactions. Although different species use diverse mechanisms for determining sex, all result in creation of sexually dimorphic gonads which produce the gametes, sperm and egg. Thus, proper sexual development of the gonad is critical to the propagation of the species. Because of that a fundamental question in reproductive biology is how sexual identity leads to sexual dimorphism.

Conservation of downstream sex determining factors, Dmrts

Sex determination mechanisms responsible for creating sexual dimorphism are diverse throughout the animal kingdom, such as cell autonomous splicing of specific sex factors in flies and hormonal secretion by gonads in mammals, but converge upon a family of conserved genes,

Dmrts (*dsx*, *mab-3* related transcription factors), that encode products that regulate sex specific gene expression. The founding member of the *Dmrt* gene family, *doublesex (dsx)*, was first identified as a mutation affecting sexual differentiation in *Drosophila* (Hildreth, 1965). The *Dmrts* play a role in sex-specific gonad development in many animal species examined, including mammals, and a human syndrome of gonad sex reversal has been linked to deletions affecting *Dmrts* (Matson and Zarkower, 2011). Thus, it appears that *dsx* plays an ancient role in the control of sexual dimorphism. The *Dmrt* family of transcription factors share a common DNA-binding domain (DM domain) but otherwise show have little sequence conservation. In animals, *Dmrt* genes are expressed tissue specifically, and integrate information about sex, time, and position in order to instruct those cells in which they are expressed in to follow along a female or male developmental program. In this chapter, I will explore how sexual identity is established in a variety of species and the role of DMRTs in sexually dimorphic gonad development.

Mammalian sex determination

In mammals, sex determination depends the presence or absence of the transcription factor Sex-determining region of the Y (Sry) located on the male-determining region of the Y chromosome (Gubbay et al., 1990; Sinclair et al., 1990). Sry encodes a DNA binding protein containing a high-mobility group (HMG) box motif (Ferrari et al., 1992). Expression of Sry in the genital ridges (gonadal primordium) of the biopotential gonad results in their development of the testes and repression of the female sex-determining pathway, whereas in the complete absence of Sry, ovary development occurs (Koopman et al., 1991). In females, Sry is sufficient to induce male development. The end result of Sry expression in the genital ridges, are sex-specific

gonads that secrete gonadal hormones (estrogen and testosterone) that initialize development of secondary sex characteristics specific to the male or female developmental programs. Similar to the sex determining gene *doublesex* (*dsx*) in flies, a fraction of direct targets of SRY have been reported in the last decade which cannot entirely account for the decision in mammalian sex determination to make testes vs. ovaries.

SRY activity in XY animals results in upregulation of its target gene, *Sox9* (SRY box containing gene 9) transcription in Sertoli cell precursors. In turn, *Sox9* expression upregulates terminal differentiation genes involved in Sertoli cell differentiation. The Sertoli cells assemble into testis cords which house the germ cells. Male-specific development of germ cells, Leydig cells which produce androgen, testis vascular cells, and non-cord cell types are induced by the Sertoli cells. The end result is the formation of a testis. *Sox9* is conserved among mammalian and non-mammalian vertebrates and is thought to be the only gene required downstream of Sry to activate the remainder of the male sex determining program (Chaboissier et al., 2004; Barrionuevo et al., 2006).

In XX animals, the absence of *Sry* initiates expression of genes involved in the female developmental program such as *Wnt4* (wingless-type MMTV integration site family, member 4) and *Foxl2* (forkhead box L2). Expression of *Wnt4* and *Foxl2* upregulate downstream genes that lead to differentiation of granulosa cells, theca cells, oocyte production, and ovarian follicles. The female developmental program is initialized in XY gonads that lack Sry function (Gilbert, 2006).

Once the testis is formed, the Sertoli cells secrete anti-Mullerian hormone (AMH) and Leydig cells secrete the steroid testosterone. Sertoli cell secretion of AMH is one of the first sexually dimorphic features of the fetal gonad. In the presence of AMH, the Mullerian duct, a

component that develops into female reproductive system is destroyed. In the presence of testosterone, tissues outside of the gonad are induced and secondary sex characteristics specific to the male develop. Specifically, testosterone stimulates the Wolffian ducts to develop into the epididymis, seminal vesicles, and vas deferens. In the absence of these masculinizing hormones, ovaries develop and thecal cells secrete estrogen which enable development of the Mullerian duct into the female reproductive system and ultimately leads to development of female secondary sex characteristics. Estrogen promotes the survival of the Mullerian duct and the absence of testosterone causes the Wolffian duct to degenerate (Gilbert, 2006).

Before the onset of sexual differentiation, genes involved in mammalian sex determination are expressed in the somatic cells of the embryonic gonad. *Dmrt1* is a gene that is expressed in the genital ridge in mammals and other vertebrates. In humans, *DMRT1* is localized on chromosome 9 (9p24.3). Hemizygosity of *Dmrt1* is associated with testicular dysgenesis and on occasion in XY humans, male to female sex reversal. Although *Dmrt1* is not required for primary sex determination in mice, *Dmrt1* is required for male gonadal differentiation in somatic cells and germ cells (Matson and Zarkower, 2011). Studies using mouse models have shown that *Dmrt1* is necessary to maintain male gonadal sex long after the foetal choice between male and female has been made. Loss of the *Dmrt1* transcription factor in Sertoli cells results in sexual cell fate reprogramming in which Sertoli cells transdifferentiate into granulosa cells. In addition, germ cells appear feminized, estrogen is produced, and other female specific cells form such as theca cells. *Dmrt1* is not only necessary but also sufficient to specify male identify in the mouse gonad. In the ovary, ectopic expression of *Dmrt1* causes sexual transdifferentiation and masculinization. *Dmrt1* expression in the ovary is sufficient to cause sexual cell fate reprogramming in which adult granolas cells transfate into Sertoli-like cells that contain male

seminiferous-like tubules (Lindeman et al., 2015; Zhao et al., 2015). These findings shed light on the etiology of human gonadal disorders as well as human syndrome linked to aberrant function of the *Dmrt1* gene.

Sex determination in birds, fish and reptiles

Like mammals, bird sex determination also occurs with the inheritance of the sex chromosome. However, instead of using XX/XY sex chromosomes, birds have defined sex chromosomes where ZW is female and ZZ is male. One or both of the avian sex chromosomes is thought to carry genes that control gonadal sex differentiation producing testes in males of ZZ chromosomal sex, and ovaries in females of ZW chromosomal sex. It is still unclear whether the mechanisms of avian sex determination depends on Z dosage or a dominant sex determining gene present on the W. Thus far, the presence of a dominant sex determining gene on the W has yet to be proven. In support of the hypothesis that avian sex determination depends on Z dosage, recent genome-wide gene expression studies in male and female heterogametic systems demonstrate an overrepresentation of sex-biased genes on the sex chromosomes (Naurin et al., 2012).

It is very clear that birds use testosterone and estrogen. However, studies on gynandromorphic birds undermined the traditional view that gonads develop into ovaries or testes that secrete sex specific hormones to masculinize or feminize the animal since cells in these animals follow their own instructions and not the gonads (Chue et al., 2011). Instead, there appears to be a cell autonomous role in which 'cells' know their sex. Taken together, sex in birds is determined on a genetic basis which result in terminal differentiation genes that mediate development of sexually dimorphic gonads and male vs. female characteristics.

The Z chromosome carries the key candidate testis determinant *DMRT1*, *dsx*- and *mab-3* related transcription factor 1. In support of the Z dosage hypothesis of avian sex determination mentioned earlier, high expression levels of *DMRT1* in genital ridges of male (ZZ) chickens triggers testicular development and expression of the Sertoli cell differentiation factor *Sox9*, whereas, a lower level of expression of *DMRT1* in females (ZW), results in ovarian development (Jakob et al., 2011; Smith et al., 2004)

Furthermore, knockdown of *DMRT1* in embryos results in feminization of male gonads. Despite there being a critical threshold of *Dmrt1* activity to determine testis fate, the role of *Dmrt1* in avian sex determination is still ambiguous, since it is unknown if *Dmrt1* is sufficient for male development. Avian sex determination appears to be similar to that of flies in which sex is determined on a genetic basis, but also requires sex specific expression of terminal differentiation genes such as *doublesex*.

The sex of most turtles and all species of crocodilians is determined by the environment after fertilization. In these reptiles, the temperature of eggs during a certain period of embryonic development determines sex (Bull, 1980). Eggs incubated at low temperatures produce one sex and eggs incubated at high temperatures produce the alternate sex. A small range of temperatures gives rise to individuals of both sexes. The temperature dependent patterns that give rise to the sex ratios vary between species (Pieau et al., 1994). Sex determination in reptiles is hormone-dependent. In reptiles, estrogen is essential for ovarian development and can override masculinizing temperatures and induce ovarian differentiation (Dorizzi et al. 1994; Rhen and Lang 1994).

The function of *Dmrt1* in temperature dependent sex determination has also been analyzed. Though the mechanisms regulating this process are unknown, there is a correlation

between increased expression of *Dmrt1* and masculinizing temperatures that promote male development (Smith et al., 1999, Kettlewell et al., 2000; Sakata et al., 2006; Chang-Soo et al., 2007).

Sex determination in Teleost fishes varies between species, but are grouped into genetic and environmental systems. Environmental factors determining sex not only include temperature, but water pH, growth rate, oxygen concentration, density, and social interactions (Baroiller et al., 2009). The genetic sex determination mechanism involves sex chromosomes that may be cytologically distinguishable (heteromorphic) or identical (homomorphic). In either case, one sex possesses two sex chromosomes and produces two types of gametes (heterogametic) and the other sex has two copies of the same sex chromosome (homogametic), producing one gamete. The male-heterogametic system follows an XX/XY system and the female-heterogametic system follows a ZZ/ZW system. Rather than being fixed, sex is a plastic phenotype in teleost fishes in which sex reversal is inducible. Unlike traditional sex determination mechanisms involving a cascade of genes acting stepwise to ultimately induce terminal genes involved in sexual differentiation, sex is a threshold phenotype triggered by genetic, parental or environmental factors that act on terminal genes and hormones (Huele, 2004).

One of the best studied sex-determining genes in fish is a *Dmrt* family member called DM domain on Y (*Dmy*; also known as *Dmrt1y*). *Dmrt1* is involved in sexual differentiation in all species of fish where it has been studied in so far: *medaka*, *ricefish*, *tilapia*, and *zebrafish*. In *medaka*, *Dmrt1* expression is testis specific and acts as a male linked dominant master regulator of sex determination that is necessary and sufficient to induce male development (Kondo, 2004; Kondo, 2006). In *tilapia*, overexpression of *Dmrt1* results in partial to complete XX female to male sex reversal. In *zebrafish*, a locus containing *Dmrt1* is implicated in being a major sex

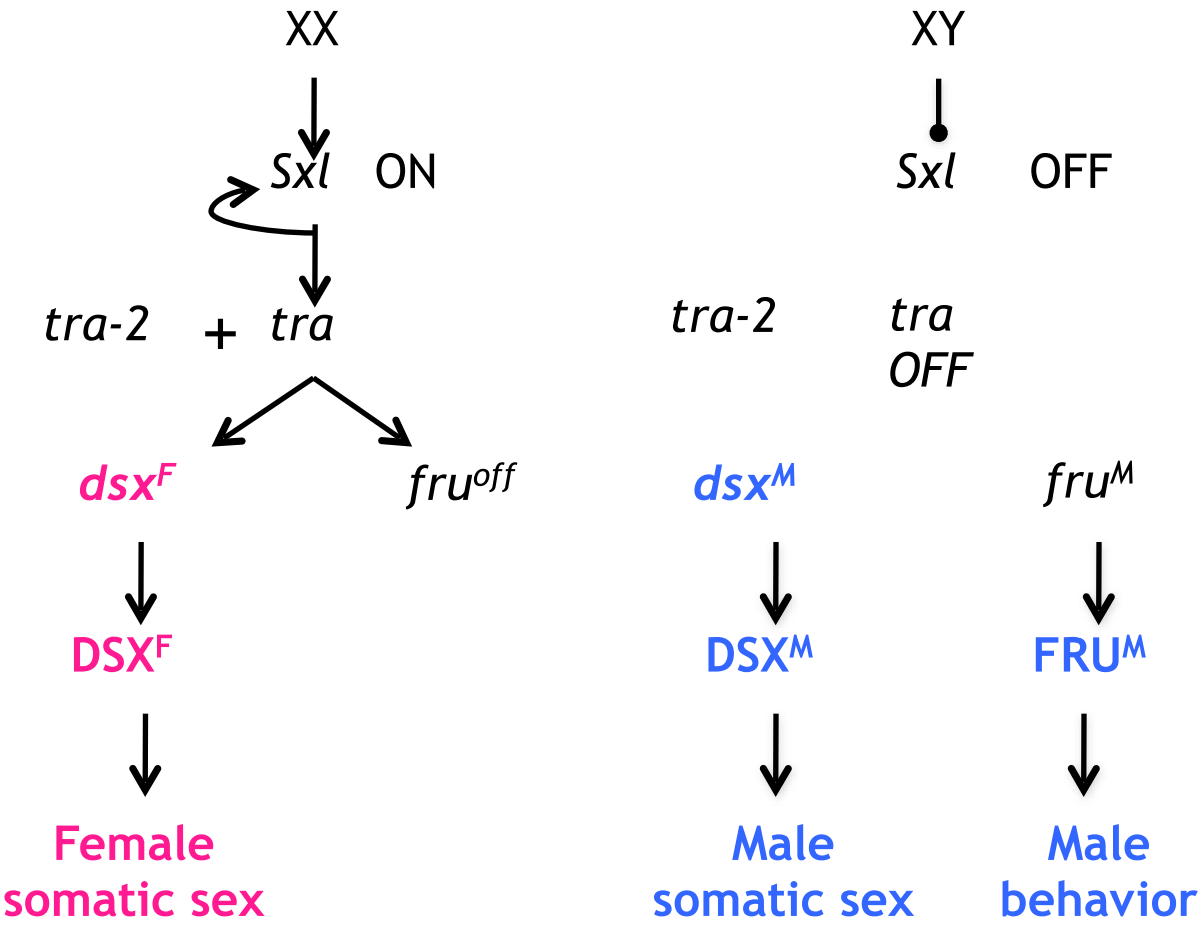
determinant. Despite the various sex-determining strategies in fishes, *Dmrt1* appears to be a conserved downstream sex regulator.

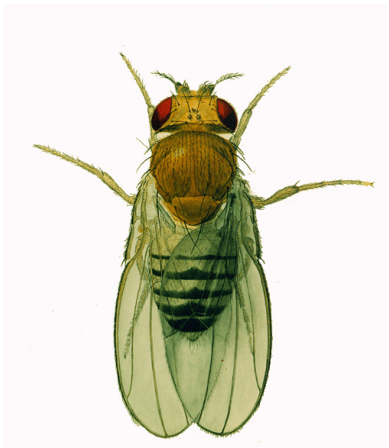
Somatic sex determination in Drosophila

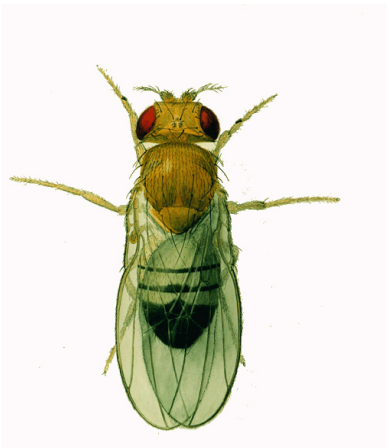
While the mechanisms by which sexual identity leads to sex-specific control of development are still being elucidated, the role of *Dmrts* is most understood in *Drosophila*. In *Drosophila*, nearly all manifestations of sexual dimorphism are regulated by the presence of sex-specific transcripts, *doublesex* (*dsx*) and *fruitless* (*fru*), generated at the end of the sex determination cascade (Figure 1.1).

Like mammals and other species discussed thus far, sex is determined on a genetic basis in *Drosophila*. However, sex determination in *Drosophila* depends on the number of X chromosomes present. Female development is initialized by having two X chromosomes, which activates the expression of *Sex-lethal* (*Sxl*). SXL is an RNA binding protein which regulates alternative splicing (Penalva and Sanchez, 2003). *Sxl* expression is initiated only in females from an early promoter (Pe) (Salz et al., 1989; Keyes et al., 1992). *Sxl* is later transcribed in both males and females from the maintenance promoter (Pm). Since only females contain functional SXL protein present from early expression, the *Sxl* transcript from the maintenance promoter can only be spliced into a functional form in the female (Bell et al., 1988). SXL expression is maintained throughout the life of the female from this auto regulatory loop (Sanchez and Nothiger 1982; Cline 1984; Bell et al., 1991).

Figure 1.1. Somatic sex determination in *Drosophila*. Sex determination is regulated by X chromosome number (XX is female, XY is male), which activates an alternative splicing cascade through female-specific expression of the RNA binding proteins Sex-lethal (SXL) and Transformer (TRA) (Fig 1). In combination with the general splicing factor Transfomer-2, TRA regulates the splicing of two known downstream target genes, *dsx* and *fruitless (fru)*. Current thinking indicates that *dsx* controls sex-specific morphology outside of the nervous system, while *dsx* and *fru* act together to control sex-specific nervous system development and behavior.


XX
Sxl ON
tra-2 + tra
dsxF
fruoff
DSXF
Female somatic sex
XY
Sxl OFF
tra-2
tra OFF
dsxM
fruM
DSXM
FRUM
Male somatic sex
Male behavior





For somatic sexual identity, SXL regulates alternative splicing of *transformer (tra)* mRNA, such that a functional TRA protein is produced only in females. In males, the absence of SXL results in splicing of *tra* into a non-functional form. TRA is an alternative splicing factor which controls sexual dimorphism (Nagoshi et al., 1988; Burtis and Baker, 1989). TRA, together with the non-sex specific splicing factor Transformer 2, acts in the alternative splicing of two downstream targets, *doublesex (dsx)* and *fruitless (fru)*, while these genes undergo default splicing in males (Belote et al., 1985; Kitadate et al., 2007). The result is the formation of different isoforms of DSX protein in males and females (DSX^M and DSX^F), while a functional FRU protein is only made in males (Chen et al., 1997; Le Bras and Van Doren, 2006). The mechanism by which these genes act to translate sexual identity into sex-specific development is still largely unknown.

The end result of the sex determination cascade are sex specific DSX isoforms that initiate development of sex specific gonads, the ovary and the testes. At this time, *dsx* is expressed specifically in the somatic gonad, while *fru* expression is not observed. This is similar to the expression pattern observed for *dsx* homologs in other animal species (the *dsx/mab3* related transcription factors, DMRTs).

Drosophila DMRT: Doublesex

Doublesex, the founding member of the DMRT gene family, was first identified as a mutation in *Drosophila* (Hildreth, 1965). Its name stems from the inter sexual phenotype of *dsx* mutants: some traits are intermediate while other tissues both male and female structures develop in parallel such as external sexual characteristics like abdominal pigmentation and the

reproductive tract. Adult *doublesex* mutant animals are sterile with gonads that do not resemble either testes or ovaries (Hildreth, 1965; Orssaud and Lauge, 1982).

The proteins encoded by *dsx* are structurally related. Both DSX isoforms share a common N terminus which contains the highly conserved DNA binding motif, the DM domain (Kopp et al., 2000; Boggs et al., 1987). The DM domain contains two zinc finger binding regions and an unstructured tail. Both the zinc finger and the tail are required for DSX binding to DNA (Narendra et al., 2002). Since both DSX isoforms share the same DNA binding domain, DSX^F homodimers can compete with DSX^M homodimers for DNA (Erdman et al., 1996).

Although both DSX isoforms share the same minimal DNA binding element, it is evident that they differ in sex-specific activity of their known downstream targets due to the differences in their C termini. There are no known co-factors with DSX^M, but DSX^F has an obligate co-factor encoded by *intersex* (*ix*). Female *intersex* mutants are phenotypically similar to *dsx* null mutants (Garrett-Engele et al., 2002). Although Intersex is present in both sexes, it only functions as a co-factor of DSX^F, which further demonstrates that sex-specific activity of DSX isoforms is conferred by differences in C termini (Chase and Baker, 1995).

Though this pathway has been established for many years, few direct targets for the DSX (or FRU) transcription factors are known *in vivo*. Direct targets of DSX include the *yolk protein* genes (*Yp1/2*) expressed in female fat body which encode yolk protein that get deposited in the oocytes; the *bric-a-brac* locus (*bab1* and *bab2*) which encode transcription factors that regulate the presence sex-specific abdominal pigmentation; *desaturase-F* (*Fad2*) which encodes a protein involved in female specific pheromone synthesis (Hempel and Oliver 2007; Hildreth, 1965;, Kitadate and Kobayashi, 2010; Randsholt and Santamaria, 2008); and *Flavin-containing monooxygenase-2* (*Fmo-2*) which is expressed in a sexually dimorphic manner but whose

function is unknown (Luo and Baker, 2015). Since interaction between DSX and its three known targets cannot account for the differences in regulation by DSX in all sexually dimorphic tissues it is expressed in, it is of great importance to identify direct and transcriptionally regulated targets of DSX.

Research on the *Drosophila* gonads, genital disc, reproductive tract, and external genitalia in males and females has shown that DSX isoforms play opposing roles on sexual differentiation: DSX^M promotes male identity and somatic development and represses female specific development by turning off target genes, whereas DSX^F induces female identity and somatic development and represses male specific development by turning on target genes.

In the case of the first known direct target of DSX, the *Fat Body Enhancer* (*FBE*) element of the *yolk protein* genes in adult fat body, DSX^M and DSX^F can bind to the same sequence in vitro but have opposing roles *in vivo*. DSX^M acts as a repressor and DSX^F acts as an activator of the *yolk protein* genes (Burtis et al., 1991; An and Wensink, 1995). In the case of the *yolk protein* genes, DSX^F blocks access of a transcriptional repressor and acts as a transcriptional activator. DSX^M functions as a repressor by preventing action of the transcriptional activator (An and Wensink, 1995). The *bric a brac* (*bab*) locus was discovered as the second target of DSX. In *Drosophila melanogaster*, male abdomens are pigmented because *bab* is repressed and in females, *bab* is activated resulting in the abdomens to remain un-pigmented. Just as the DSX isoforms act on yolk protein genes in the fat body, DSX^M acts as the repressor and DSX^F acts as the activator in the the abdomen (Williams et al., 2008). Interestingly, the third known direct target of DSX, *desatF*, is only mediated by one DSX isoform, DSX^F. In adult female oenocytes, DSX^F binds and directly activates *desatF* expression. Strikingly, there is no repressive function for DSX^M in regulating *desatF* expression since loss of DSX function in males does not lead to

an up regulation of *desatF* in oenocyte cells (Shirangi, et al., 2009). In the case of the most recently identified DSX target, *Fmo-2*, DSX^M acts as the repressor and DSX^F acts as the activator of *Fmo-2* expression in the midgut (Luo and Baker, 2015).

Examples of genes whose expression is DSX dependent yet not directly regulated by DSX demonstrate that DSX^M functions as an activator and DSX^F functions as a repressor (Waterbury et al., 1999; Dauwalder et al., 2002; Arbeitman et al., 2004; Goldman and Arbeitman, 2007). However, the current logic is that sex specific isoforms of DSX act on its known target genes such that DSX^F is an activator and DSX^M is a repressor. This logic may not be universal for how DSX regulates all of its target genes and in all tissues DSX is expressed in. Thus, further work on putative targets of DSX are needed to understand the regulatory modes of each isoform.

The sex determination hierarchy acts at a particular time and place to control development and maintenance of sexually dimorphic tissues. We know that direct and indirect target genes that respond to *dsx* expression are different in each cell type in which DSX functions (Christiansen et al., 2002). There are two classes of genes potentially regulated by DSX: Early versus constituitively regulated genes (Burtis et al., 1991). Early genes are genes regulated by DSX at a particular time during development for a specific tissue to form and may not be expressed at later stages of development after initiation of that tissue. Constituitively regulated genes are genes that are regulated in the adult fly for maintenance of a particular tissue. As a consequence of genes required during development and maintenance of sexual dimorphism, it is likely that DSX regulates different targets in different tissues.

While the genes involved in sex determination, such as *Sxl*, are expressed in all somatic cells (Bopp et al., 1991), terminal genes in the hierarchy, *dsx* and *fru*, are not. Temporally, the highly tissue-specific pattern of *dsx* expression begins as early as mid embryogenesis in the

gonad, some imaginal tissues do not express *dsx* until pupal stages, and some tissues require *dsx* expression to be turned off (Hempel and Oliver, 2007; Lee et al., 2002; Rideout et al., 2010; Robinett et al., 2010). Spatially, *dsx* may be expressed throughout a tissue or only in a few cells in others. The dynamic tissue specific pattern of *dsx* expression illustrates that not all cells of the animal "know" their sex. In order for a cell to translate its sex chromosome constitution into information it can use to control sex-specific development, it needs to express a transcription factor, such as DSX or FRU, which is regulated by the sex chromosome genotype.

Formation of the Drosophila gonad

In most animals, the somatic support cells that will house and nurture the germ cells during gametogenesis are formed from the mesoderm during development. In *Drosophila*, the somatic gonadal precursors (SGPs) are mesodermal cells specified in bilateral clusters within three abdominal parasegments (10-12) that will form on either side of the embryo ("parasegments" (PS) are the units of segmental identity along the A/P axis). Each mesodermal parasegment is divided into an anterior ("*even skipped (eve)* domain") and posterior ("*sloppy paired* domain") and the SGPs form within the eve domain, while in other PS this domain gives rise to the fat body (Riechmann et al., 1998; Moore et al., 1998). The dorsal-ventral axis is also divided into distinct domains, and the SGPs in PS 10-12 form within the dorso-lateral domain, below the dorsal domain that gives rise to the heart and visceral mesoderm. The transcription factors Eyes Absent (EYA) and Zn Finger Homeodomain 1 (ZFH-1) are expressed in this domain and are critical for SGP and fat body specification (reviewed by Jemc, 2011).

An additional cluster of SGPs forms in PS13, but is located more ventrally in the mesoderm and does not depend on the genes that form the dorso-lateral domain (DeFalco et al.,

2003). However, these cells do share characteristics of SGP identity, such as expression of *eya*, Six4 and *dsx* (DeFalco et al., 2003; Clark et al., 2006; Hempel and Oliver, 2007). These are known as "male-specific SGPs" (msSGPs) since they will survive only in males. Thus, each gonad in males is initially formed from 4 separate clusters of SGPs.

The homeotic genes (HOX), *abd-A* and *Abd-B*, promote SGP formation in parasegments and give rise to the unique identity of SGP clusters. *abd-A* specifies anterior SGP identity and *Abd-B* specifies msSGPs. By stage 15, anterior-posterior identity of the gonad is established and can be observed by different markers (e.g., anterior SGPs express escargot while posterior SGPs express higher levels of *eya* and *Wnt2*), indicating their distinct identities.

The earliest manifestations of sexual dimorphism in the gonad are observed at the time of gonad formation, when sex-specific gene expression is observed in SGPs (DeFalco et al., 2008; Casper and Van Doren, 2009). In males, SGPs signal to the male germ cells through the JAK/STAT pathway (Wawersik et al., 2005), and in females the msSGPs undergo apoptosis (DeFalco et al., 2003). At this time, *dsx* is expressed specifically in SGPs (Hempel and Oliver, 2007; DeFalco et al., 2008). Consistent with this, *dsx* controls all aspects of initial sexual dimorphism in gonad development so far described.

Testis Development

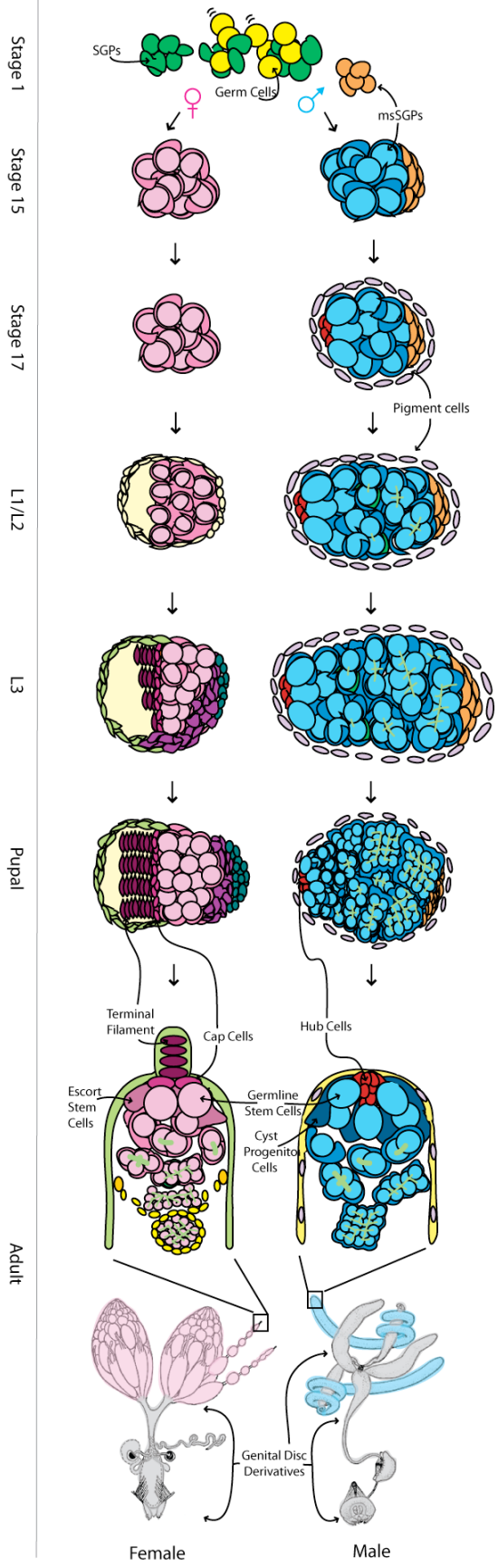
The embryonic gonad has a number of distinct cell types, including germ cells, anterior SGPs, posterior SGPs and msSGPs. Further, these cells have a clear sexual identity as indicated by the expression of a number of sex-specific molecular markers. In the next stage of gonad development, this information is then combined to allow sexually dimorphic formation of the ovaries and testes. One interesting difference between testis and ovary development is that the

testis is largely complete by the end of embryogenesis or the early larval stages (24-30 hours after fertilization), while morphogenesis of the ovary occurs later during the larval-pupal transition (5 days after fertilization) (Figure 1.2).

The male-specific SGPs

The presence of the msSGPs at the posterior of the male embryonic gonad is one of the first sexual dimorphisms to be identified in the developing gonad. msSGPs are present in males but not females. In addition to expression of molecular markers characteristic of SGPs, the msSGPs express the transcription factor SOX100B (DeFalco et al., 2003). Remarkably, SOX100B is the fly homolog of mammalian Sox9, a protein that is essential for sex determination in the gonads of mouse and human (Jakob et al., 2009). *Sox100B* mutants exhibit severe defects in testis development but show no defects in the ovary (Nanda et al., 2009). Thus, this transcription factor is critical for gonad sexual dimorphism in diverse species. The msSGPs ultimately give rise to the terminal epithelium of the testis, which is critical for the final differentiation of spermatids, and likely also plays a role in connecting the testis to other portions of the reproductive tract.

Figure 1.2. Development of sexual dimorphism in the *Drosophila* gonad. Embryonic stages (1, 15, 17) are as described (Campos-Ortega and Hartenstein, 1985). L1/L2, 1st and 2nd instar larvae. L3, 3rd instar larvae. The adult stage depicts the apical end of a single ovariole in the female and testis in the male. The gonad begins as a bi-potential organ (Stage 1). Germ cells and somatic gonadal precursors (SGPs) interact and form an embryonic gonad by stage 15. msSGPs are initially specified in both sexes, but join only the posterior of the male gonad. In the male, pigment cells and the embryonic hub form during stage 17. The female gonad undergoes ovary morphogenesis during late L3 to make individual ovarioles and establish the female germline stem cell niche at the L3-pupal transition. In adults, germline stem cells (GSCs) contact the somatic niche formed by cap cells in females and hub cells in males. Somatic stem cells (cyst progenitor cells in male and escort stem cells in females) also contact the niche. germ cells (yellow at st 12, then light pink or blue); SGPs (green, st 12, then pink or blue; msSGPs (orange); hub cells (red); pigment cells (gray), terminal filaments (dark purple); cap cells (magenta); escort stem cells (purple); hub (red), cyst progenitor cells (navy blue). Adult ovaries (pink), testes (blue), and genital disk derivatives (gray).


Stage 1
Stage 15
Stage 17
L1/L2
L3
Pupal
Adult
SGPs
Germ Cells
msSGPs
Pigment cells
Terminal Filament
Cap Cells
Hub Cells
Escort Stem Cells
Germline Stem Cells
Cyst Progenitor Cells
Genital Disc Derivatives
Female
Male

While msSGPs are found only in the male gonad, these cells are initially present in both sexes prior to gonad formation but undergo sex-specific apoptosis in females (DeFalco et al., 2003). In males, these cells actively migrate to join the SGPs and germ cells in forming the gonad (Clark et al., 2007). If programmed cell death is blocked in the msSGPs, then they survive in females and join the gonad, indicating that sex-specific apoptosis is the key step controlling their sexually dimorphic development (DeFalco et al., 2003). Interestingly, even though these cells express DSX (Hempel, 2007; DeFalco et al., 2008), they do not decide for themselves whether to behave as male or female, but receive their instructions from the other SGPs. Experiments with sexual mosaics indicate that the behavior of the msSGPs depends on the sex of the SGPs, rather than on the sex of the msSGPs. Further, in the absence of SGPs, the msSGPs die in both sexes, indicating that a sex-specific survival cue from male SGPs is what determines the fate of the msSGPs (DeFalco et al., 2008).

The behavior of the msSGPs highlights a general principle of how sexual dimorphism is created in the gonad; many cell types do not decide for themselves what sex they should be, but rather receive signals from other cell types, namely the SGPs, about what sexual path they should follow. This is true for the msSGPs and is also true for the pigment cell precursors and the germ cells. This is contrary to the commonly-held belief that, in *Drosophila*, "all cells decide for themselves" what sex they should be (Wolpert, 2006). Thus, the creation of sexual dimorphism in the fly gonad is similar to that of gonad development in mammalian systems, where non-autonomous sex determination is known to be common.

The pigment cell precursors

An additional male-specific cell type arises in the testis during the last stage of embryogenesis (stage 17, ≈20 hours after fertilization). These cells were originally discovered because they also express SOX100B and were thought to be derived from the msSGPs, but were found to be of independent origin (DeFalco et al., 2008). They surround the outside of the embryonic gonad and give rise to the pigment cells that ensheath the adult testis and parts of the reproductive tract. While the function of these cells is not clear, it has been speculated that they may influence testis function by means such as hormonal signaling.

Interestingly, the pigment cell precursors provide another example of non-autonomous sex determination in the developing gonad. These cells are not part of the embryonic gonad at the time of its formation, but are specified from the surrounding fat body via male-specific signaling from the SGPs (DeFalco et al., 2008). SGPs express *Wnt2* in a male-specific manner under control of the somatic sex determination pathway, and *Wnt2* is necessary and sufficient for formation of pigment cell precursors(DeFalco et al., 2008). The sex of the fat body itself does not influence this process; both male and female fat body are competent to produce pigment cell precursors when provided with the *Wnt2* signal (DeFalco et al., 2008). Further, male-specific pigment cell precursor formation is under control of *dsx* (DeFalco et al., 2008), which is expressed in SGPs but not the pigment cell precursors themselves (Hempel, 2008). Instead, *dsx* regulates *Wnt2* expression in the SGPs (DeFalco et al., 2008). The sex-specific recruitment of neighboring cell types to become part of the developing gonad is a common feature of gonad development in different animals, such as in the recruitment of mesonephric cells into the mouse testis, a process that is controlled by the sex of the gonad rather than the sex of the mesonephros (reviewed by Brennan et al., 2004).

The testis stem cell niche

One of the most interesting aspects of testis development is the formation of the germline stem cells (GSCs) and the somatic environment, or "niche", that regulates them. GSCs are critical for the production of large numbers of gametes throughout an extended period of adult life. In *Drosophila*, both the testis and the ovary have GSCs as well as somatic stem cells that produce the differentiated somatic cells that nurture the germline during gametogenesis. However, there is clear sexual dimorphism in the male and female GSCs, somatic stem cells, and the niches that control these stem cells. In other species, such as mouse and human, there is a clear GSC population only in the testis, and the ovary has a more limited capacity to produce gametes. How sexual dimorphism in both the soma and the germline lead to differences in the gonad stem cell systems is thus an important question for understanding sex-specific development and reproductive health.

The testis stem cell system is located at the proximal end of the testis, and is organized around an important cluster of somatic cells that make up the "hub" (Figure 1.2). The GSCs are arranged around the hub, along with the Cyst Stem Cells (CySCs), the somatic stem cells that produce the cyst cells that support the germline during spermatogenesis. The hub acts as a key organizing center by adhering to the stem cells, allowing them to remain in the niche and regulating their pattern of cell division. The hub also signals to the stem cells through multiple signaling pathways, including JAK/STAT, TGF-ß and other pathways (reviewed in Matunis et al., 2012; Zoller and Schulz, 2012). In addition to producing the cyst cells, the CySCs also form an important part of the niche that regulates the GSCs (Leatherman, 2008)

The hub forms during the last stage of embryogenesis (Stage 17, ≈20-24 hours after fertilization) from a subset of SGPs in the anterior region of the embryonic gonad (Le Bras and

Van Doren, 2006). Initially, SGPs with a combination of anterior and male identities express unique molecular markers such as *escargot*. Subsequently, a subset of these anterior male cells coalesce to form a tightly associated cluster of cells that express many of the molecular markers also expressed in the adult hub. This includes expression of several cell-cell adhesion molecules (DE-cadherin, DN-cadherin, Fasciclin-3) that are likely to mediate the sorting of these cells away from other SGPs and into the tight cluster of the embryonic hub (Le Bras and Van Doren, 2006). A subset of germ cells associate with the embryonic hub as it forms and take on GSC identity. Although all germ cells in the male gonad initially exhibit activation of the JAK/STAT pathway, this response becomes restricted to those germ cells associated with the hub (Sheng et al., 2009). Further, the hub-proximal germ cells exhibit the oriented divisions characteristic of GSCs and produce progeny that enter differentiation (Sheng et al., 2009). Thus, by the embryo/larval transition, a fully functional stem cell niche has formed in the testis.

Since differentiating spermatogonial cysts are observed soon after formation of the hub and GSCs, it is likely that the CySCs also form at this time. However, no specific markers are available that distinguish these cells at this early time point. Lineage analysis indicates that the CySCs are derived from the same pool of anterior SGPs that form the hub (Dinardo et al., 2011). It is also likely that the posterior SGPs associate with those germ cells that are not selected to become GSCs to directly form spermatogonial cysts ("one shot spermatogenesis") (Sheng et al., 2009). However, it remains possible that posterior SGPs use their unique identity to contribute to some other cell type in the adult testis that remains to be identified.

If the same pool of anterior SGPs (PS10 and PS11) give rise to hub cells and CySCs, how then is the hub vs. CySC decision made such that the correct number of hub cells form? Recent work indicates that the transcription factor Bowl is important for this decision (Dinardo et al.,

2011). Fewer hub cells are formed in *bowl* mutants, while removing an inhibitor of *bowl*, known as *lines*, increases hub cell number. Thus, *bowl* appears to act as a "pro-hub" factor in the hub vs. CySC decision. Interestingly, when *bowl* is activated by loss of *lines* in adult CySCs, they take on some hub cell character, indicating that the hub cells and CySCs remain closely related and their distinction must be maintained even in the adult testis (Dinardo et al., 2011).

It is also clear that cell-cell signaling is important for determining which SGPs will form hub cells. Signaling between somatic cells via the Notch pathway regulates hub cell number, with *Notch* mutants showing a strong decrease in hub cells (Kitadate and Kobayashi, 2010; Okegbe et al., 2011). Notch is expressed broadly in SGPs, and evidence for Notch activity is seen in SGPs in both anterior and posterior regions of the gonad (Kitadate and Kobayashi, 2010; Okegbe et al., 2011). Thus, it is likely that anterior SGPs that have activated Notch will form hub cells, while those that do not form CySCs. Since Notch is known to upregulate *bowl* in other contexts, one possibility is that Notch acts at least in part through *bowl* to specify hub cell identity (Dinardo et al., 2011). However, there is disagreement about the nature and location of the ligands for the Notch pathway that control hub formation. Kitadake *et al*. observe that Serrate is expressed in SGPs and is the major ligand affecting hub formation (Dinardo et al., 2011). However, Okebe and DiNardo report that the Delta ligand plays the major role, with a lesser role for Serrate, but that neither is expressed in the gonad (Okegbe et al., 2011). Instead, they postulate that Delta expression in the neighboring midgut signals to the SGPs. Further research is needed to resolve which somatic cells are important for activating Notch signaling to control hub cell specification.

Signaling from the germ cells to the SGPs through receptor tyrosine kinases (RTKs) is also involved in hub cell specification (Kitadate and Kobayashi, 2010; Kitadate et al., 2007).

Posterior SGPs express the Sevenless (SEV) RTK, while all SGPs express the Epidermal Growth Factor Receptor (EGFR). Loss of either or both of these RTKs increases the number of SGPs that take on hub identity, but there is still a clear bias toward formation of hub cells from anterior SGPs, indicating that the anterior/posterior patterning of the SGPs is the major determinant specifying which SGPs can form hub cells. Ligands for each of these RTKs (BOSS and Spitz) are expressed in germ cells, suggesting that germ cells limit the production of an important component of their own stem cell niche, the hub. One intriguing idea is that this allows the system to compensate for situations where fewer germ cells reach the gonad (Kitadate and Kobayashi, 2010). When germ cells are plentiful, they restrict the number of cells that take on hub identity and allow for sufficient CySC production. However, when few germ cells are present, hub cell number is increased to ensure that some germ cells contact hub cells and become germline stem cells. Indeed, when germ cell number is reduced experimentally, this mechanism appears to be important for increasing the likelihood of germline stem cell production and formation of a functional testis (Kitadate and Kobayashi, 2010).

One problem with this model is that, if the same pool of cells generates both hub cells and CySCs, when the number of hub cells is increased, CySC number would be decreased. As CySCs and cyst cells are also essential for gametogenesis, it would be counterproductive to increase hub cells at the expense of CySCs unless some other mechanism was in place to ensure CySC production. It is important to note, however, that manipulation of either the Notch or RTK pathways alters hub cell number by only a modest amount. For example, double mutants of both *sev* and *Egfr* increase hub cell number from about 9 to about 16 cells from a total of approximately 40 SGPs (Kitadate and Kobayashi, 2010). Thus, these pathways can alter the likelihood that an SGP will choose hub cell identity over CySC identity, but they are unlikely to

convert the entire pool of anterior SGPs to hub cells, for the obvious reason that loss of CySCs would be disastrous for testis function.

Later steps in testis development

Since the testis contains a functioning stem cell system and begins producing differentiating spermatogonial cysts by the early larval period, little needs to change regarding basic testis function during the nine days of development between larval and adult stages. However, some mutations affect only the testis at the larval/pupal transition (e.g. *Sox100B)* (Nanda et al., 2009), indicating there may be additional changes or checkpoints that occur at this important developmental transition. This is the time when ovary morphogenesis begins, and the developmental "timer" that controls ovary development may also trigger changes to the testis.

One important event that happens during the early pupal period is attachment of the testis to the rest of the reproductive tract, which develops from the genital imaginal disc (reviewed by Camara et al., 2008). The genital disc derivatives attach to the posterior of the testis, at the site where the msSGPs form the terminal epithelium. The genital disc is also the source of muscle cells that migrate to form a uniform layer around the testis underneath the pigment cell layer (Kozopas et al., 1998). Attachment of the genital disc derivatives and the migration of the muscle cells causes the testis to elongate from the spherical shape of the larval testis to the extended testis "tube" of the adult.

Ovary morphogenesis and niche formation

The adult Drosophila ovary is divided into 16-20 units called ovarioles that continually produce eggs. Each ovariole contains 6-8 egg chambers, each of which arises from a germ stem

cell. Germ stem cells reside in the germanium located at the anterior tip of each ovariole. On average, the germanium hosts 2-3 germ stem cells adjacent to the niche which include somatic cells such as 5-7 cap cells attached to 8-10 terminal filament cells. Niche cells produce the ligand Decapentaplegic (Dpp, a BMP2/4 homologue) which maintains germ stem cells attached to the cap cells. Dpp signaling in the germ stem cells is required to repress *bag of marbles* (*bam*), a major differentiation gene. The GSC undergoes asymmetric cell division such that the daughter cell remains at the niche as a GSC and the other, called a cystoblast, is removed from the niche and enters oogenesis. Continuous egg production depends on the function of germ stem cells and their close association to the somatic cells (terminal filaments, cap cells) that direct their renewal and differentiation (escort cells).

While much is known about the function and molecular mechanisms employed by the adult germ stem cell niche, the mechanisms involved in organizing the stem cell units in the *Drosophila* ovary are less clear. During larval development the somatic gonadal precursors and primordial germ cell lineages proliferate and differentiate. The driving force in the first two days of larval ovary development is mainly proliferation. The somatic gonad remains morphologically undifferentiated until later larval stages. Within the next 24 hours (mid-third instar), anterior cells or swarm cells migrate laterally past a cluster of germ cells to form part of the posterior of the ovary (Courderc et al., 2002). Once at the posterior of the ovary, these somatic cells differentiate to form the intermolecular stalk, basal stalk, and basal cells during the later larval and pupal stages (Courderc et al., 2002). During these stages, germ cells closely interact with the somatic cells, coined intermingled cells, a group of interstitial cells that encapsulate the germ cells throughout their development. A subset of somatic cells express *Engrailed* that remain at the anterior of the ovary become terminal filament cells and cap cells (Bolivar et al., 2006 and Godt

and Laski, 1995). Development of terminal filaments also depends on other Tramtrack-group nuclear factors, such as *Bric-a-brac* (*bab*). The Bab protein is expressed in terminal filament cells before they differentiate into disc-shaped cells (Godt and Laski, 1995).

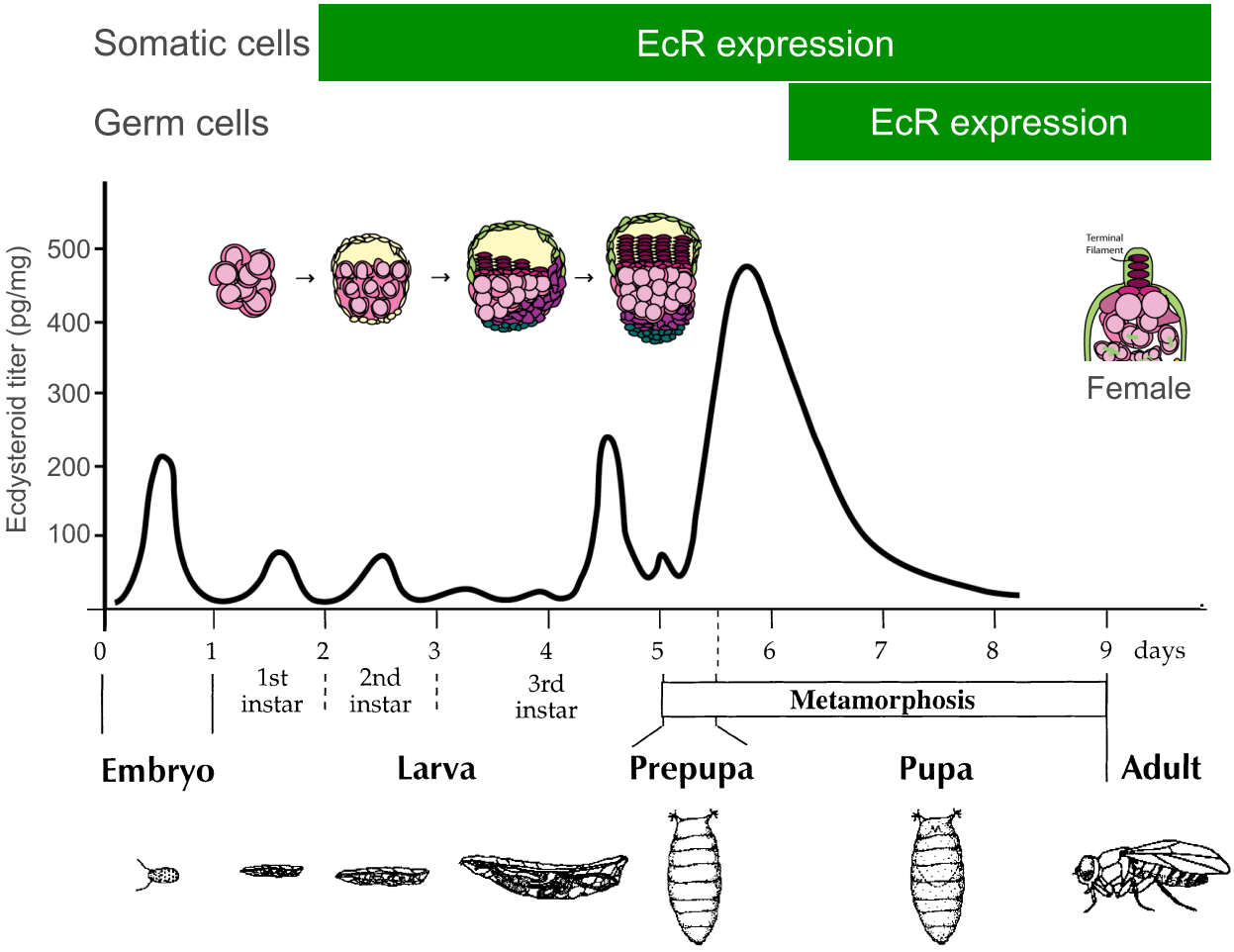
The first somatic niche component to form is the terminal filament lineage. Terminal filament cells are specified just prior to third instar. At mid third instar, terminal filament differentiation initiates and terminal filament lineage specification continues throughout the late larval period. It is not until 96 hours after egg laying that the disc shaped terminal filament cells appear. Terminal filaments form progressively from the medial to the lateral side of the ovary by intercalation of terminal filament cells (Godt and Laski, 1995). During the last 24 hours of larval development, the terminal filament cells stack to form 16-20 terminal filaments and are post-mitotic.

Importantly, around 96 hours there is a shift of emphasis from proliferation to differentiation for both somatic and germ cells. During the larval to pupal transition, terminal filament cells recruit intermingled cells through notch signaling to promote cap cell formation as the base of the terminal filament stacks (Xie et al., 2008?). Once terminal filaments and cap cells form, anterior most primordial germ cells that associate with the cap cells via E-Cadherin to become adult germ stem cells (Song et al., 2002). PGCs that do not associate with cap cells differentiate into the first germ line cysts and egg chambers (Zhu and Xie, 2003). By late pupal stages, anterior somatic ovary cells migrate between terminal filament stacks to separate each individual ovariole.

The developmental "timer" that controls timing and progression of ovarian morphogenesis is controlled by the steroid hormone ecdysone signaling pathway (Figure 1.3). This finding is not surprising since the vast changes that occur during insect metamorphosis of

larval tissues into adult structures are all mediated by ecdysone hormone signaling. There are two distinct phases of ovarian morphogenesis. The first phase consists of proliferation and formation of terminal filament stacks and the second is characterized by the onset of germ cell differentiation and morphogenesis of somatic cell types in the ovary. The transition between the first and second phases is mediated by changes in the levels of the hormone ecdysone. Ecdysone is synthesized and secreted in pulses by the prothoracic gland. EcR can bind to the ecdysone hormone alone. However, in order for EcR to bind to ecdysone responsive elements and activate transcription of its targets, EcR requires a co-factor Ultraspiracle (USP). In the absence of ecdysone hormone, the presence of EcR results in repression of target genes (Thummel 1996, 1990, 1995; Koelle et al. 1991). During early third instar, EcR and USP are expressed and function to repress ecdysone targets and prevent differentiation of germ cell and somatic cells. The purpose of this step is to provide the germ cells time to proliferate and increase the size of the ovary (Gancz et al., 2003). During third instar, the first phase of ovarian morphogenesis involving the period of growth is initialized via an ecdysone pulse that triggers

Figure 1.3. Summary of *EcR* expression patterns in the ovary. Ovarian morphogenesis in *Drosophila* follow a strict temporal script. *EcR* expression is observed in the somatic cells of the gonad throughout larval development of the ovary and persist in the adult. *EcR* is not expressed in germ cells until pupal stages and also persists in the adult. The female gonad undergoes ovary morphogenesis during late L3 to make individual ovarioles and establish the female germline stem cell niche at the L3-pupal transition. Development correlates with pulses of ecdysone hormone. germ cells (light pink); SGPs (pink), terminal filaments (dark purple); cap cells (magenta); escort stem cells (purple), basal cells (green).


Somatic cells
EcR expression
Germ cells
EcR expression
Ecdysteroid titer (pg/mg)
500
400
300
200
100
Terminal Filament
Female
0
1
2
3
4
5
6
7
8
9
days
1st instar
2nd instar
3rd instar
Metamorphosis
Embryo
Larva
Prepupa
Pupa
Adult

terminal filament proliferation and germ cell differentiation a few hours later. These important events are not observed at first and second instar which correlate with the earlier (lower) peaks of ecdysone. The second phase occurs over the next 48 hours, during the late larval to pupal transition, characterized by the onset of germ cell differentiation and morphogenesis of somatic cell types. These events correlate with a robust increase in ecdysteroid level and with expression of early ecdysteroid induced response genes such as the Tramtrack-group nuclear factor, Broad-Complex (BR-C)(Andres et al. 1993). EcR and its co-factor, USP, are highly expressed in all somatic cell types just prior to differentiation. The importance of ecdysone signaling is further illustrated when flies mutant for *EcR* and *usp* produce defects in ovarian morphogenesis including a delay in the onset of terminal filament differentiation, delayed competition of terminal filament formation, and delayed ovarian differentiation (Hodin and Riddiford, 1998).

Role of doublesex in sex-specific niche formation

In *Drosophila*, both the ovary and testis have germline stem cells that are controlled and maintained by their niches (Figure 1.2). Critical components of these niches are the "hub" in males and "terminal filaments and cap cells in females (Gönczy, 1992; Le Bras, 2006) (Figure 1.2). The hub, terminal filaments and cap cells have related functions in organizing and signaling to the gonad stem cells, and also have a common origin from somatic cells of the embryonic gonad. However, they are also very different between males and females in several key ways. The hub develops in males at the end of embryogenesis (22-24 hours after egg laying, AEL) while the terminal filaments and cap cells form in females during late larval to pupal stages (4-5 days AEL) (King, 1970; Zhu, 2003). A single hub forms per testis from ≈8-10 cells, while 16-18 sets of terminal filaments and cap cells form (one for each ovariole in the ovary) from a much

larger pool of progenitor cells. Morphologically, the hub is a highly compacted ball of cells while the terminal filament is an elongated chain of stacked-disc shaped somatic cells. There are key differences in how the hubs vs. terminal filaments signal to the gonad stem cells. Although the two niches are quite different morphologically, there are similarities in how they act to nurture the germline stem cells (Dansereau, 2008; Gilboa, 2004; Lin, 2002; Fuller, 2007).

In the Van Doren lab, we have investigated the role that *dsx* plays in sexually dimorphic gonad formation and identified many interesting aspects on how sexual identity influences the early somatic gonadal cells to produce such distinct stem cell niches (Camara et al., unpublished). Since *dsx* is normally required for proper male and female development, most tissues take on an intersexual morphology when mutant for *dsx*. However, for the gonad stem cell niches we found a strikingly different result. In the late embryo of *dsx* mutant animals, we found that sex-specific cell types described above appears fully male-like in *dsx* mutants; both XY and XX *dsx* mutant embryonic gonads have msSGPs, pigment cell precursors and hubs. These hubs formed at the correct developmental time, and were able to adhere and signal normally with the GSCs. In late larvae, at the time when terminal filaments normally form, we found that in 50% of animals the hubs were lost and terminal filaments formed in their place. This occurred in 50% of both XY and XX animals. In the animals that "switched" from hubs to terminal filaments, we observed that the hub cells reentered the cell cycle and appeared to contribute to the larger pool of cells that are needed to form terminal filaments. The terminal filaments that formed interacted normally with the GSCs. Terminal filaments are known to induce cap cell formation and, indeed, we observed cap cell formation whenever terminal filaments were present. In both XX and XY adults, 50% of the gonads still had hubs and 50% had terminal filaments. These modes of development appeared to be mutually exclusive—

evidence was seen for hubs or terminal filaments in a gonad, but not both. However, for gonads in the same animal, one could have a hub and the other terminal filaments. If a hub was allowed to form in the presence of DSX^M, but then switched to DSX^F at 2nd instar (using a $tra2^{ts}$ allele), the hubs no longer switched to terminal filaments. Thus, hubs formed under the influence of DSX^M were more resistant to "switching" to TF than were *dsx* mutant hubs. From this data, we make the following conclusions and build our model (Figure 1.4) for how *dsx* regulates sexual dimorphism in the gonad stem cell niche. Since hubs and terminal filaments can form in the complete absence of *dsx* function, *dsx* is not required for the niche to form. However, they now do so stochastically, independent of the sex chromosome genotype. Thus, *dsx* does not appear to be instructive for niche formation, but instead acts to ensure that the proper pathway (hub vs. terminal filaments) is activated in the proper sexual genotype (XY vs. XX). We propose that endogenous hub and terminal filaments pathways exist that can be modulated by *dsx*, but do not require *dsx* for their function. Since both XX and XY gonads form hubs in *dsx* mutants, *dsx* is required in females to block hub formation at the time that the hub normally forms in the embryo. The niches are more plastic than previously thought. Even though a hub forms in *dsx* mutants, it is still sensitive to feminizing influence and can change fate and form terminal filaments. This is similar to the switch from male to female in the mouse gonad when DMRT1 activity is removed. In *dsx* mutant larvae, 50% of both XX and XY gonads switch from hub to terminal filaments. Thus, *dsx* is normally required in females to ensure a robust response to the "pro terminal filaments" pathway so that all female gonads form terminal filaments. Further, *dsx* is normally required in males to repress the "pro terminal filaments" pathway and ensure that all male gonads maintain hub fate. The cells of the gonad must be able to "agree" with each other about which pathway to activate, since one gonad usually makes either a hub or terminal

filaments, but not both. Thus, we propose that there is a non-autonomous signal by which the cells of the gonad agree on their sexual fate.

Conclusions

Since interaction between DSX and its four known targets can not account for the differences in regulation by DSX in all sexually dimorphic tissues it is expressed in, we sought to identify direct and transcriptionally regulated targets of DSX. In Chapter 2, we undertook multiple experimental approaches that allowed us to identify genes that were bound by DSX, genes whose expression changed in response to DSX perturbation, and genes that function in *dsx*-expressing cells. In Chapter 3, we performed a detailed analysis on RNA-seq experiments to identify genes that are DSX dependent. The gonad represents an excellent model to understand *doublesex* at a particular time and place. Thus, we used our extensive genomics data along with developmental biology of gonad stem cell niches to identify ecdysone signaling as a target of DSX in our model of sex-specific gonad niche development. In Chapter 4, we examine the *Ecdysone receptor* gene, which is involved in ecdysteroid signaling, for roles in sex-specific gonad development.

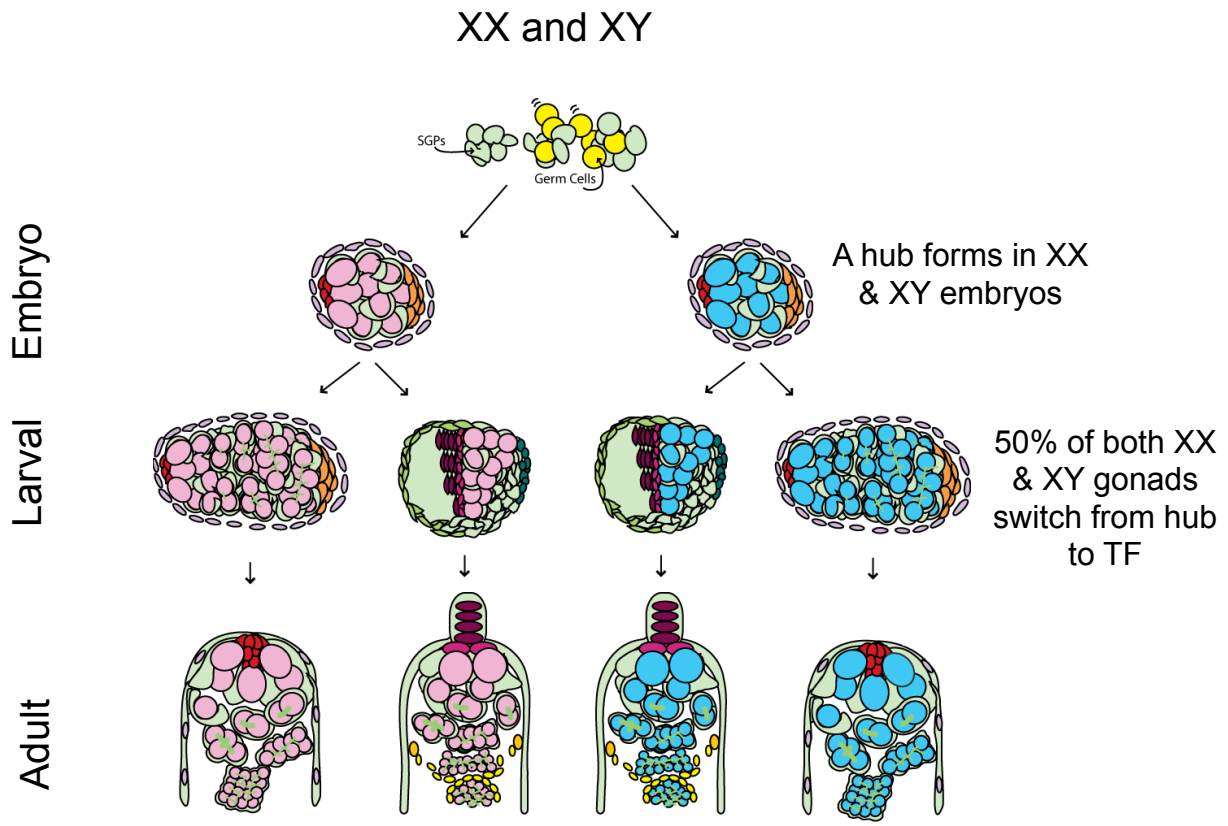


Figure 1.4. Role of *doublesex* in sexually dimorphic niche formation. In the late embryo of *dsx* mutant animals, both XY and XX *dsx* mutant embryonic gonads have male specific somatic gonadal precursors (msSGPs), pigment cell precursors and hubs. In late larvae, at the time when terminal filaments normally form, 50% of animals with hubs are lost and terminal filaments form in their place. In adults, 50% of the gonads still have hubs and 50% have terminal filaments.

CHAPTER 2: SEX- AND TISSUE-SPECIFIC FUNCTIONS OF DROSOPHILA DOUBLESEX TRANSCRIPTION FACTOR TARGET GENES

This Chapter was performed in collaboration with the scientists listed below and is published:

Sex- and tissue-specific functions of Drosophila doublesex transcription factor target genes. Dev Cell. 2014 Dec 22;31(6):761-73. doi: 10.1016/j.devcel.2014.11.021.

Emily Clough[1§], Erin Jimenez[2§], Yoo-Ah Kim[3§], Cale Whitworth[1,2,§,*], Megan C. Neville[4], Leonie Hempel[1], Hania Pavlou[4], Zhen-Xia Chen[1], David Sturgill[1], Ryan Dale[5], Harold E. Smith[6], Teresa M. Przytycka[3#], Stephen F. Goodwin[4#], Mark Van Doren[2#] and Brian Oliver[1#] [1]Section of Developmental Genomics, Laboratory of Cellular and Developmental Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, 50 South Drive, Bethesda, MD 20892, USA. [2]Department of Biology, Mudd Hall, Johns Hopkins University, Baltimore, MD 21218, USA. [3]Computational Biology Branch, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, 8600 Rockville Pike, Bethesda, MD 20814, USA. [4]Department of Physiology, Anatomy and Genetics, Sherrington Building, Parks Road, University of Oxford, Oxford OX1 3PT, UK. [5]Office of the Chief, Laboratory of Cellular and Developmental Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, 50 South Drive, Bethesda, MD 20892, USA. [6]Genomics Core, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, 8 Center Drive, Bethesda, MD 20892, USA. [§]Equal contributors listed alphabetically[#]Senior Authors*Contact: cale.whitworth@nih.gov; tel. 301-435-8405; fax none.

Supporting datasets not incorporated into this thesis are available through: doi: 10.1016/j.devcel. 2014.11.021.

Author Contributions

E.C., C.W., L.H., M.N., and H.P performed DSX molecular biology and occupancy experiments. C.W., E.J., E.C., and D.S. performed RNA-seq. C.W., E.J., and E.C. performed molecular genetics experiments. Y.-A.K., Z.-X.C, R.D., and E.C performed conservation and meta-analysis. R.D. and D.S. performed bioinformatics analysis. H.S. sequenced samples. E.C. was consortium coordinator. C.W, B.O., and M.V.D. wrote the manuscript. T.P., S.G., M.V.D., and B.O. supervised the project.

Summary

Primary sex determination "switches" evolve rapidly in animals, but Doublesex (DSX) related transcription factors (DMRTs) act downstream of these switches to control sexual development in most animal species. Drosophila *dsx* encodes female- and male-specific isoforms (DSX^F and DSX^M), but little is known about how *dsx* controls sexual development, whether DSX^F and DSX^M bind the same or different targets, or how DSX proteins direct different outcomes in diverse tissues. We undertook genome-wide analyses to identify DSX targets using *in vivo* occupancy, binding site prediction, and evolutionary conservation among 20 Drosophila species. We find that DSX^F and DSX^M bind thousands of the same targets in multiple tissues in both sexes, yet these targets have sex- and tissue-specific functions. Interestingly, the DSX targets show considerable overlap with targets identified for mouse DMRT1. DSX targets include transcription factors and signaling pathway components providing for direct and indirect regulation of sex-biased expression.

Introduction

Genetically encoded sexual dimorphism allows males and females to differ in appearance, physiology, and behavior. Differences in sperm and egg morphology and the delivery systems that ensure that they meet are often obvious, but there are also many subtle aspects of sex differentiation impacting organs and physiology throughout the body. Controlling the sexual development of a broad range of cell types is a challenge since sex-biased gene expression advantageous in one tissue may be detrimental in another. Sex determination systems must therefore provide organism-level, sex-specific modulation of gene expression that is simultaneously compatible with a range of tissue-specific requirements. This suggests that sex-specific and tissue-specific gene expression must be tightly integrated. How this occurs is not well understood.

Primary sex determination signals vary in animals, but the *doublesex* and *mab-3* related transcription factors (DMRTs) are known to control sex determination and differentiation in diverse species (Zarkower, 2013). For example, XY humans with deletions of 3 DMRT genes exhibit sex reversal (Raymond et al., 1999). The founding member of the DMRT family, *doublesex* (*dsx*), is required for the majority of sexually dimorphic morphology in *Drosophila* and also acts in the nervous system to regulate behavior (reviewed by Camara et al., 2008). A female-specific alternative splicing cascade results in the production of functional Transformer (TRA) and Transformer 2 (TRA2), which regulate sex-specific alternative splicing of *dsx*. In females, the *dsx* pre-mRNA is spliced to encode DSX^F protein. In the absence of TRA, male-specific splicing of *dsx* pre-mRNA occurs, and this transcript encodes DSX^M (Burtis and Baker, 1989; Nagoshi et al., 1988). The DSX^F and DSX^M isoforms have the same DNA-binding and dimerization domains but have different C-termini (Bayrer et al., 2005; Zhang et al., 2006).

Intersex (IX) binds specifically to the C-terminus of DSX^F and is required for DSX^F function (Yang et al., 2008), suggesting that the sex-specific C-termini are effector domains interacting with co-factors to modulate gene expression.

In the absence of *dsx*, flies show intersexuality at the cellular level; therefore, *dsx* is not required for the production of sex-specific structures, but, rather, mediates which sex-specific structure is formed. The sex-specific DSX isoforms have opposite genetic activities and either the absence of all DSX or the presence of both DSX^F and DSX^M isoforms results in similar phenotypes. For example, the dsx^D allele can only produce DSX^M such that XX; dsx^D/+ animals produce both DSX^F (from the wild-type allele) and DSX^M (from the dsx^D allele) resulting in an intersexual phenotype resembling that shown by *dsx* loss-of-function (Nagoshi and Baker, 1990).

In addition to regulation by alternative splicing, *dsx* is expressed in a highly tissue-specific manner indicating that cells are on a "need to know" basis with respect to sex (Hempel and Oliver, 2007; Lee et al., 2002; Rideout et al., 2010; Robinett et al., 2010). The *dsx* locus is expressed in subsets of neurons, gut cells, somatic cells of the gonad, adipose, and hepatic tissues. These cell types derive from all the primary germ layers (ectoderm, endoderm, mesoderm) and have diverse roles in metabolism, gametogenesis, morphology, and behavior. While the transcriptional inputs to *dsx* expression are not fully understood, Drosophila HOX genes and other patterning genes regulate *dsx* in at least some tissues (Foronda et al., 2012; Tanaka et al., 2011; Wang et al., 2011; Wang and Yoder, 2012; Yoder, 2012).

Although DSX was identified in 1965 (Hildreth, 1965) and cloned in 1988 (Baker and Wolfner, 1988), there are still few defined DSX targets and these cannot explain the full array of sexually dimorphic morphologies and behaviors regulated by *dsx* in *D. melanogaster*. Most of the known DSX target genes were identified on a case-by-case basis (Burtis et al., 1991; Shirangi

et al., 2009; Williams et al., 2008). While there have been large numbers of genome-wide expression studies on the sexes (Samson and Rabinow, 2013), there have been fewer attempts to link this expression directly to DSX (Lebo et al., 2009). One study identified genes with sex-biased expression in genital discs and showed that expression of these genes was *dsx*-dependent, but did not address whether these were directly or indirectly regulated (Chatterjee et al., 2011). DSX^F occupancy has been examined genome-wide and was filtered using a precise 13-mer predicting the presence of 23 direct target genes (Luo et al., 2011), but this analysis does not capture the known DSX targets and is therefore unlikely to represent all targets in the genome. We have combined an extensive DSX occupancy study on both DSX^F and DSX^M isoforms in multiple tissues with comparative genomic analyses of 20 species of *Drosophila* and expression profiling of a tissue during an acute switch in DSX isoform. We have tested the functional relevance of these targets using an unbiased *dsx* genetic interaction screen, and determined the roles of predicted DSX targets in *dsx*-expressing cells. Our analysis reveals that DSX is bound to many of the same targets in males and females and in different tissues, indicating that DSX action at these genes is regulated downstream of DSX binding. Further, we find a striking conservation of DSX targets with orthologous genes identified as targets of mouse DMRT1 (Murphy et al., 2010), suggesting that the control of sexual dimorphism may be similar in diverse animal species.

Materials and methods

Fly stocks

Fly stocks were obtained from the Bloomington Drosophila Stock Center (Cook et al., 2010), the Transgenic RNAi Project (Ni et al., 2011), and from the B.S. Baker lab and other

generous members of the Drosophila community. See FlyBase for gene and allele descriptions (Marygold et al., 2013) for *tra2*^{ts2} (FBal0017028), *tra2*^{ts1} (FBal0017027), *P{UAS-tra.F}20J7* (FBti0010566), *P{tubP-GAL80ts}7* (FBti0027798), *P{tubP*-GAL4*}LL7* (FBti0012687), *dsx*^{D} (FBal0003200), GAL4$^{dsx.KI}$ (FBal0277019), dsx^{GAL4} (FBal0244772), *gpp*^{X} (FBal0175658), lilli^{A17-2} (FBal0103689), *w*^{1118}, and *Oregon R*. Bloomington Deficiency Kit II stocks used can be found in Table S3. Alleles tested for genetic interaction with *dsx* and TRiP RNAi lines can be found in Table S4. Information on *UAS-Dam-myc* (*Dam*), *UAS-Dam-myc-dsx*^{F} (*Dam-dsx*^{F}), and *UAS-Dam-myc-dsx*^{M} (*Dam-dsx*^{M}) can be found below. Flies were grown on standard Bloomington Drosophila Stock Center (Bloomington, IN, USA) or Drosophila Species Stock Center (San Diego, CA, USA) medium at 25°C unless otherwise noted. Information about unspecified fly stocks can be found at http://flybase.bio.indiana.edu.

Table 2.1. Deficiency Kit Stocks used in *dsx^{D}* screen

Deficiency Name	Flybase ID	right breakpoint*	left breakpoint*	Deleted Segment*	shifted dsxD phenotype
Df(2R)14H10W-35	FBab0029927	14020748	13625546	54E5--55B7	No
Df(2R)M41A4	FBab0001993	469718	348419	41A--41A	No
Df(2R)M60E	FBab0001997	20878249	20828423	60E6--60E11	No
Df(2R)Kr10	FBab0001981	21321410	20842993	60E10--60F5	No
Df(2R)Px2	FBab0002027	20492825	20155881	60C6--60D9	No
Df(2R)59AD	FBab0022239	19172912	18563151	59A1--59D4	No
Df(2R)X58-12	FBab0022257	18573320	18080991	58D1--59A1	No
Df(2R)AA21	FBab0004927	17282113	16976985	57B19--57E6	No
Df(2R)P34	FBab0002011	15172984	14551832	55E6--56C1	No
Df(2R)Jp8	FBab0001974	12212018	11946618	52E1--53C1	No
Df(2R)CX1	FBab0022242	9950134	8587926	49C1--50D5	No
Df(2R)vg-C	FBab0002234	8827958	8408228	49B2--49E2	No
Df(2R)CB21	FBab0024848	8418392	7955746	48E--49A	No
Df(2R)en30	FBab0002176	7796537	7467536	48A3--48C8	No
Df(2R)X1	FBab0002048	6224893	5809508	46C2--47A1	No
Df(2R)B5	FBab0010225	5876110	5525043	46A1--46C12	No
Df(2R)H3C1	FBab0024375	4547749	3749018	43F--44D8	No
Df(2R)ST1	FBab0002037	3749018	2539440	42B3--43E18	No
Df(2L)TW161	FBab0001648	21871823	19868511	38A6--40B1	No
Df(2L)pr-A16	FBab0001868	20687000	18881927	37B2--38D5	No
Df(2L)cact-255rv64	FBab0022223	17817472	16348738	35F6--36D	No
Df(2L)TE35BC-24	FBab0001606	16125490	14648531	35B4--35E2	No
Df(2L)b87e25	FBab0001747	15092774	13428842	34C1--35C1	No
Df(2L)N22-14	FBab0022196	8399332	8399212	29C1--30C9	No
Df(2L)TE29Aa-11	FBab0001569	8404869	8083671	28E4--29C1	No
Df(2L)Dwee1-W05	FBab0026801	6939661	6880269	27C2--27C5	No

Df(2L)E110	FBab0001456	6514155	5682521	25F3--26D11	No
Df(2L)cl-h3	FBab0001758	6110195	5240239	25D2--26B5	No
Df(2L)ed1	FBab0001792	4073079	3619317	24A2--24D4	No
Df(2L)JS17	FBab0022191	3298921	2875792	23C1--23E2	No
Df(2L)C144	FBab0022184	2955279	2517598	22F4--23C3	No
Df(2L)dpp[d14]	FBab0001786	2721176	2425049	22F1--23A2	No
Df(2L)dp-79b	FBab0001773	2425019	1428804	22A2--22E1	No
Df(2L)ast2	FBab0001695	1794258	626409	21D1--22B3	No
Df(2R)ED4065	FBab0035310	20830362	20290189	60C8--60E8	No
Df(2R)BSC155	FBab0044848	20145426	19968928	60B8--60C4	No
Df(2R)BSC161	FBab0044965	13372333	13192288	54B2--54B17	No
Df(2R)Exel7131	FBab0038038	10247930	10118170	50E4--50F6	No
Df(2R)BSC134	FBab0044840	10153306	10063603	50E1--50E6	No
Df(2R)Exel7130	FBab0038037	10100288	9960585	50D4--50E4	No
Df(2R)BSC132	FBab0044814	5748332	5482319	45F6--46B4	No
Df(2L)BSC151	FBab0044888	22139023	21828581	40A5--40E5	No
Df(2L)Exel6049	FBab0037887	22019296	21828252	40A5--40D3	No
Df(2L)BSC147	FBab0044819	13665417	13445419	34C1--34C6	No
Df(2L)BSC159	FBab0045063	13536086	13290762	34B4--34C4	No
Df(2L)BSC145	FBab0044836	11001945	10967405	32C1--32C1	No
Df(2L)BSC143	FBab0044911	10333704	10209408	31B1--31D9	No
Df(2L)ED611	FBab0032338	8419818	8382851	29B4--29C3	No
Df(2L)BSC111	FBab0040373	8362842	8240266	28F5--29B1	No
Df(2L)BSC142	FBab0044910	8012787	7774037	28C3--28D3	No
Df(2L)BSC110	FBab0040372	5064620	5029595	25C1--25C4	No
Df(2L)ED250	FBab0031954	4821294	4477085	24F4--25A7	No
Df(2L)BSC106	FBab0038756	417947	291728	21B7--21C2	No
Df(2L)C'	FBab0001429	21754733	21731512	h35--40A1	No
Df(2L)BSC5	FBab0029534	6448047	5980164	26B1--26D2	No
Df(2L)BSC41	FBab0037635	8048798	7495902	28A4--28D9	No
Df(2R)BSC44	FBab0037744	13326201	13115252	54B1--54B10	No
Df(2R)14H10Y-53	FBab0029942	13637373	13489244	54D1--54E7	No

Df(2R)BSC26	FBab0029945	15357320	15108494	56C4--56D10	No
Df(2L)BSC4	FBab0029533	456219	267408	21B7--21C3	No
Df(2L)BSC16	FBab0029789	464653	456209	21C3--21C8	No
Df(2L)BSC37	FBab0037634	2458422	2132199	22D1--22F2	No
Df(2L)BSC28	FBab0029867	3298921	3030443	23C5--23E2	No
Df(2L)BSC31	FBab0029974	3486273	3344927	23E5--23F5	No
Df(2L)drm-P2	FBab0029744	3636539	3443841	23F3--24A2	No
Df(2L)BSC7	FBab0029660	6880269	6499666	26D10--27C1	No
Df(2L)Trf-C6R31	FBab0022203	8110561	8071809	28DE;28DE	No
Df(2L)BSC17	FBab0029725	9951990	9659196	30C3--30F1	No
Df(2L)BSC32	FBab0031351	11067015	10557961	32A1--32D1	No
Df(2L)BSC36	FBab0031352	11067015	10557961	32D1--32E1	No
Df(2L)BSC30	FBab0029975	13393105	12942602	34A3--34B9	No
Df(2R)BSC29	FBab0029979	5501106	5285324	45D3--45F6	No
Df(2R)BSC39	FBab0037639	7992013	7736387	48C5--48E1	No
Df(2R)BSC40	FBab0037640	8118269	7955746	48E1--48E10	No
Df(2R)BSC3	FBab0029423	8469798	8119980	48E12--49B6	No
Df(2R)BSC11	FBab0029712	11063783	10133068	50E6--51E4	No
Df(2R)BSC49	FBab0037746	13326201	12675592	53D9--54B10	No
Df(2R)robl-c	FBab0028920	13437865	13369175	54B17;54C4	No
Df(2R)k10408	FBab0028916	13369175	13335515	54B16;54B16	No
Df(2R)BSC45	FBab0037745	13637373	13448649	54C8--54E7	No
Df(2R)BSC22	FBab0029834	16163236	15329256	56D7--56F12	No
Df(2R)vir130	FBab0024859	19371530	18654199	59B;59E1	No
Df(2L)net-PMF	FBab0001854	324485	1	21A1--21B8	Yes
Df(2L)BSC109	FBab0038758	5145500	5073453	25C4--25C8	Yes
Df(2L)Exel6011	FBab0037853	5305646	5147258	25C8--25D5	Yes
Df(2L)spd[j2]	FBab0024845	7369686	6798469	27B2--27F2	Yes

Df(2L)XE-3801	FBab0024841	7829302	7117728	27E2--28D1	Yes
Df(2L)BSC50	FBab0037839	10280557	9973195	30F5--31B1	Yes
Df(2L)Prl	FBab0001522	12626238	11437223	32F1--33F2	Yes
Df(2L)r10	FBab0001903	16637042	15271253	35D1--36A7	Yes
Df(2L)TW137	FBab0001645	19035083	17154051	36C2--37B10	Yes
In(2R)bw[VDe2L]Cy[R]	FBab0005116	1816712	209942	41A-42A3	Yes
Df(2R)H3E1	FBab0024377	4881316	4366981	44D1--45A1	Yes
Df(2R)Np5	FBab0024011	5411775	4848845	44F12--45E3	Yes
Df(2R)w45-30n	FBab0002275	5412274	4993940	45A6--45E3	Yes
Df(2R)en-A	FBab0002173	7531762	7042690	47D7--48B2	Yes
Df(2R)BSC18	FBab0029741	10023306	9882236	50D1--50D7	Yes
Df(2R)Jp1	FBab0001967	12098689	10683767	51C3--52F9	Yes
Df(2R)PC4	FBab0002014	14670420	13778264	55A1--55F2	Yes
Df(2R)Exel7162	FBab0038054	16201140	16132691	56F11--56F16	Yes
Df(2R)BSC19	FBab0029795	16420290	16163226	56F12--57A4	Yes

*Breakpoint and cytology data were obtained from Flybase. When available, exact breakpoints were used. For deficiencies with unknown exact breakpoints, the estimated breakpoints provided by Flybase (Dmel release 5) were used.

Table 2.2. Alleles tested for genetic interaction with *dsx^D*

Flybase Allele ID	Allele name	Gene Flybase ID	Gene Name
FBal0000085	abd-A[P10]	FBgn0000014	abd-A
FBal0033317	Abd-B[M5]	FBgn0000015	Abd-B
FBal0000622	aop[1]	FBgn0000097	aop
FBti0005165	P{PZ}bun[00255]	FBgn0259176	bun
FBal0137288	chm[14]	FBgn0028387	chm
FBal0003003	dpp[d6]	FBgn0000490	dpp
FBal0028155	ds[33K]	FBgn0000497	ds
FBal0003530	Egfr[f2]	FBgn0003731	Egfr
FBal0260468	fz2[MI02902]	FBgn0016797	fz2
FBal0175658	gpp[x]	FBgn0264495	gpp
FBti0004966	Gug[J5A3]	FBgn0010825	Gug
FBal0005292	H[1]	FBgn0001169	H
FBal0094021	InR[E19]	FBgn0013984	InR
FBal0103689	lilli[A17-2]	FBgn0041111	lilli
FBal0102516	lola[e76]	FBgn0005630	lola
FBal0064630	Mad[k00237]	FBgn0011648	Mad
FBal0012016	mam[8]	FBgn0002643	mam
FBti0114837	P{EP}mgl[G17430]	FBgn0261260	mgl
FBal0012950	neur[11]	FBgn0002932	neur
FBal0035437	pnt[Δ88]	FBgn0003118	pnt
FBti0004231	P{PZ}salm[03602]	FBgn0261648	salm
FBti0021310	P{GT1}Smr[BG01648]	FBgn0265523	Smr
FBal0034034	Su(Tpl)[10]	FBgn0014037	Su(Tpl)
FBti0005919	P{lacW}tai[k15101]	FBgn0041092	tai
FBal0031315	tj[PL3]	FBgn0000964	tj
FBal0016825	tkv[8]	FBgn0003716	tkv
FBal0123565	wb[BG02232]	FBgn0261563	wb
FBal0146678	wts[3-17]	FBgn0011739	wts
FBti0005416	P{PZ}zfh1[00865]	FBgn0004606	zfh1

Table 2.3. RNAi lines used to test putative DSX targets

Genotype	RNAi Flybase ID	Gene	Gene Flybase ID
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.GLV21008}attP2	FBst0035644	abd-A	FBgn0000014
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.GLV21012}attP2	FBst0035647	Abd-B	FBgn0000015
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01977}attP2	FBst0039057	Alh	FBgn0261238
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01256}attP2	FBst0034909	aop	FBgn0000097
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS02207}attP2	FBst0041673	ap	FBgn0000099
y[1] v[1] P{y[+t7.7] v[+t1.8]=TRiP.JF02273}attP2	FBst0026731	Atg1	FBgn0260945
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.GLV21072}attP2/ TM3, Sb[1]	FBst0035707	bab1	FBgn0004870
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.GLV21085}attP2	FBst0035720	bab2	FBgn0025525
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02419}attP2	FBst0027074	bowl	FBgn0004893
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02585}attP2	FBst0027272	br	FBgn0000210
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02954}attP2	FBst0028322	Bunched	FBgn0259176
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.GLV21048}attP2	FBst0035683	Camta	FBgn0259234
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00563}attP2	FBst0034603	CG11009	FBgn0036318

y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS01322}attP2	FBst0034334	CG14476	FBgn0027588
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS01274}attP2/ TM3, Sb[1]	FBst0034925	CG1894	FBgn0039585
y[1] v[1] P{y[+t7.7] v[+t1.8]=TRiP.JF02953}attP2	FBst0028321	CG30089	FBgn0050089
y[1] v[1] P{y[+t7.7] v[+t1.8]=TRiP.JF02711}attP2	FBst0027557	CG30382	FBgn0050382
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02348}attP2	FBst0027027	chameau	FBgn0028387
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.GLV21035}attP2/ TM3, Sb[1]	FBst0035670	cindr	FBgn0027598
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS01628}attP40	FBst0036737	Cyp4g1	FBgn0010019
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS02208}attP2	FBst0041674	domino	FBgn0020306
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02455}attP2	FBst0036779	dpp	FBgn0000490
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00759}attP2	FBst0032964	ds	FBgn0000497
P{KK111266}VIE-260B	FBti0142441	dsx	FBgn0000504
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS00107}attP2	FBst0034798	ear	FBgn0026441
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02384}attP2	FBst0036773	egfr	FBgn0003731
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02515}attP2	FBst0029353	Eip74EF	FBgn0000567
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01530} attP40	FBst0035780	Eip75B	FBgn0000568
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.GLV21073}attP2	FBst0035708	fal	FBgn0028380

y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02722}attP2	FBst0027568	fz2	FBgn0016797
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00160}attP2	FBst0034842	gpp	FBgn0264495
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00756}attP2	FBst0032961	Grunge	FBgn0010825
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01182}attP2	FBst0034703	H	FBgn0001169
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.GLV21039}attP2	FBst0035674	hang	FBgn0026575
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.GLV21049}attP2	FBst0035684	Hsc70-4	FBgn0001219
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF01183}attP2	FBst0031594	InR	FBgn0013984
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.GLV21092}attP2	FBst0037545	ix	FBgn0001276
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS01990}attP2	FBst0039070	jbug	FBgn0028371
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.GLV21020}attP2	FBst0035655	Jheh2	FBgn0034405
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS02335}attP40	FBst0041938	jub	FBgn0030530
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01063}attP2	FBst0034589	lea	FBgn0002543
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HM05155}attP2	FBst0028944	lid	FBgn0031759
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS00958}attP2	FBst0033995	ligatin	FBgn0041588
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01066}attP2	FBst0034592	lilli	FBgn0041111

y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.GLV21086}attP2	FBst0035721	lola	FBgn0005630
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.GLV21013}attP2/ TM3, Sb[1]	FBst0035648	mad	FBgn0011648
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02881}attP2	FBst0028046	mam	FBgn0002643
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.GLV21091}attP2/ TM3, Sb[1]	FBst0037486	mamo	FBgn0264981
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02485}attP2	FBst0029324	mgl	FBgn0261260
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02048}attP2	FBst0026023	neur	FBgn0002932
y[1] v[1] P{y[+t7.7] v[+t1.8]=TRiP.JF01299}attP2	FBst0031341	Nurf-38	FBgn0016687
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.GLV21077}attP2	FBst0035712	Pdcd4	FBgn0030520
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS01286}attP2	FBst0034611	Pde11	FBgn0085370
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01259}attP2	FBst0034911	plod	FBgn0036147
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02227}attP2	FBst0031936	pnt	FBgn0003118
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02596}attP2	FBst0027284	Prosap	FBgn0040752
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01843}attP40/ CyO	FBst0038374	Psn	FBgn0019947
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS01386}attP2/ TM3, Sb[1]	FBst0034392	puc	FBgn0243512
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01029}attP2	FBst0034557	pyd3	FBgn0037513

y[1] v[1] P{y[+t7.7] v[+t1.8]=TRiP.JF01766}attP2	FBst0031250	RapGAP1	FBgn0264895
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS00224}attP2	FBst0033353	RunxA	FBgn0083981
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02361}attP2/TM3, Sb[1]	FBst0027035	rutabaga	FBgn0003301
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS00594}attP2	FBst0033714	salm	FBgn0261648
y[1] v[1] P{y[+t7.7] v[+t1.8]=TRiP.JF02375}attP2	FBst0027049	sbb	FBgn0010575
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS00581}attP2	FBst0033704	Set1	FBgn0040022
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS00871}attP2	FBst0033925	skpC	FBgn0026175
y[1] v[1] P{y[+t7.7] v[+t1.8]=TRiP.JF01228}attP2	FBst0031467	sli	FBgn0264089
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02413}attP2	FBst0027068	smr	FBgn0265523
y[1] v[1] P{y[+t7.7] v[+t1.8]=TRiP.JF02225}attP2	FBst0031934	Snoo	FBgn0085450
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01450}attP2	FBst0035036	Socs36E	FBgn0041184
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS00277}attP2	FBst0033399	Su(Tpl)	FBgn0014037
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HM05182}attP2	FBst0028971	tai	FBgn0041092
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.GLV21046}attP2	FBst0035681	term	FBgn0003683
y[1] sc[*] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS01069}attP2	FBst0034595	tj	FBgn0000964

y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.HMS02185}attP40	FBst0040937	tkv	FBgn0003716
y[1] sc[*] v[1] P{y[+t7.7] v[+t1.8]=TRiP.HMS01316}attP2	FBst0034329	tll	FBgn0003720
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF03238}attP2	FBst0029559	wb	FBgn0261563
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02741}attP2	FBst0027662	wts	FBgn0011739
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF03119}attP2/TM3, Sb[1]	FBst0031965	yki	FBgn0034970
y[1] v[1]; P{y[+t7.7] v[+t1.8]=TRiP.JF02509}attP2/TM3, Sb[1]	FBst0029347	zfh1	FBgn0004606

Table 2.4. Genetic interaction and RNAi data for putative DSX target genes

FlyBase Gene Name	FlyBase Identifier	Occupancy and DSX sites	Phenotype Descriptions
abd-A	FBgn0000014	*abd-A* is occupied by DSX and is a member of occupancy cluster 2 (Fig 2A). The strength of DSX binding at *abd-A* is in the 95th percentile (gene-level PWM score=0.931) of all genes. The conservation of DSX binding at *abd-A* is in the 88th percentile (conservation index score=30.48) of all Dmel genes.	Wildtype *abd-A* promotes female genitalia and ovary development. Driving a shRNAi against *abdA* (FBst0035644, TRiP GLV21008) with *dsx*-Gal4 (FBal0277019) results in female-specific adult genitalia and gonad defects. Males are unaffected. The ovaries of adult *abd-A* knockdown flies are not organized into distinguishable ovarioles with successive stages of egg development. No mature eggs with dorsal appendages are present. The terminal filament cells are not stacked properly and thus lack wildtype organization. The female genitalia fail to evert during morphogenesis and are recessed into the abdomen. *abd-A* genetically interacts with *dsx* to promote female-like gonad development. 44% of *XX; abd-A^{P10}/dsx^{D}* gonads have hubs while 56% have terminal filaments (n=39). This is in contrast to the phenotype of XX; *dsx^{D}*/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).

Abd-B	FBgn0000015	*Abd-B* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Abd-B* is in the 98th percentile (gene-level PWM score=0.981) of all genes. The conservation of DSX binding at *Abd-B* is in the 96th percentile (conservation index score=62.14) of all Dmel genes.	Wildtype *Abd-B* promotes gonad development in males and genitalia development in both sexes. Driving a shRNAi against *Abd-B* (FBst0035647, TRiP GLV21012) with *dsx*-Gal4 (FBal0277019) results in testes that have a bulbous anterior testis tip as well as an accumulation of motile sperm at the most distal portion of the testis. The male genitalia exhibit a rotation defect relative to the wildtype orientation. The female genitalia have a vaginal plate that does not open to the exterior of the fly and is missing vaginal teeth on the most ventral portion of the plate. As a result, these flies are not able to lay eggs even though normal eggs are present in the abdomen. *Abd-B* genetically interacts with *dsx* to promote male-like gonad development. 2% of *XX; Abd-B^{M5}/dsx^{D}* gonads have hubs while 98% have terminal filaments (n=41). This is in contrast to the phenotype of XX; *dsx^{D}*/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).
Alh	FBgn0261238	*Alh* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Alh* is in the 96th percentile (gene-level PWM score=0.941) of all genes. The conservation of DSX binding at *Alh* could not be determined due to the lack of similar sequences in other Drosophila species.	Driving a shRNAi against *Alh* (FBst0039057, TRiP HMS01977) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

aop	FBgn0000097	*aop* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *aop* is in the 92nd percentile (gene-level PWM score=0.798) of all genes. The conservation of DSX binding at *aop* is in the 83rd percentile (conservation index score=21.55) of all Dmel genes.	Driving a shRNAi against *aop* (FBst0034909, TRiP HMS01256) with *dsx*-Gal4 (FBal0277019) results in late-pupal lethality. No such lethality was observed in *dsx*-Gal4 alone or in Oregon R controls. *aop* genetically interacts with dsx to promote female-like gonad development. 32% of XX; *aop*^{1}/+; dsx^{D}/+ gonads have hubs while 68% have terminal filaments (n=47). This is in contrast to the phenotype of XX; *dsx*^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).
ap	FBgn0000099	*ap* is occupied by DSX and is a member of occupancy cluster 3 (Fig 2A). The strength of DSX binding at *ap* is in the 93rd percentile (gene-level PWM score=0.840) of all genes. The conservation of DSX binding at *ap* is in the 94th percentile (conservation index score=46.92) of all Dmel genes.	Driving a shRNAi against *ap* (FBst0041673, TRiP HMS02207) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

Atg1	FBgn0260945	*Atg1* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Atg1* is in the 96th percentile (gene-level PWM score=0.963) of all genes. The conservation of DSX binding at *Atg1* is in the 78th percentile (conservation index score=17.00) of all Dmel genes.	Wildtype *Atg1* promotes male gonad development and genitalia development in both males and females. Driving a shRNAi against *Atg1* (FBst0026731, TRiP JF02273) with *dsx*-Gal4 (FBal0277019) results in a sex-specific adult male gonad phenotype and non-sex-specific adult male and female genitalia phenotypes. The germline stem cell niche (i.e. hub cells, cyst stem cells, and germline stem cells) of these adult male *Atg1* knockdown flies are often "pinched off" and protruding from the apical testis tip. No similar phenotypes were observed in dsx-Gal4 alone or in Oregon R controls. The genitalia of male *Atg1* knockdown flies are recessed into the abdomen of these flies and are not as large as normal male genitalia. Knockdown female genitalia are recessed into the abdominal cavity of the fly and are reduced in overall size although all components are present.
bab1	FBgn0004870	*bab1* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *bab1* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *bab1* is in the 97th percentile (conservation index score=68.89) of all Dmel genes.	Driving a shRNAi against *bab1* (FBst0035707, TRiPGLV21072) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

bab2	FBgn0025525	*bab2* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *bab2* is in the 97th percentile (gene-level PWM score=0.969) of all genes. The conservation of DSX binding at *bab2* is in the 95th percentile (conservation index score=54.59) of all Dmel genes.	Driving a shRNAi against *bab2* (FBst0035720, TRiP GLV21085) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
bowl	FBgn0004893	*bowl* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *bowl* is in the 87th percentile (gene-level PWM score=0.620) of all genes. The conservation of DSX binding at *bowl* is in the 89th percentile (conservation index score=32.42) of all Dmel genes.	Driving a shRNAi against *bowl* (FBst0027074, TRiP JF02419) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
br	FBgn0000210	*br* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *br* is in the 98th percentile (gene-level PWM score=0.996) of all genes. The conservation of DSX binding at *br* is in the 97th percentile (conservation index score=74.21) of all Dmel genes.	Driving a shRNAi against *br* (FBst0027272, TRiP JF02585) with *dsx*-Gal4 (FBal0277019) results in late-pupal lethality. No such lethality was observed in *dsx*-Gal4 alone or in Oregon R controls.

bun	FBgn0259176	*bun* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *bun* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *bun* is in the 34th percentile (conservation index score=4.962) of all Dmel genes.	Wildtype *bun* promotes male gonad and genitalia development. Driving a shRNAi against bun (FBst0028322, TRiP JF02954) with *dsx*-Gal4 (FBal0277019), results in male-specific gonad and genitalia phenotypes. Adult *bun* knockdown male gonads resemble 3rd instar testes being rounded and not elongated like normal adult testes. These testes are connected to the genital tract. The external genitalia of the adult *bun* knockdown males are missing key male structures including the penis apparatus, clasper teeth, and lateral lobes. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. Adult female bun knockdown flies are unaffected. *bun* genetically interacts with *dsx* to promote female-like gonad development. 37% of XX; *P{PZ}bun*^{00255}/+; *dsx*^{D}/+ gonads have hubs while 63% have terminal filaments (n=35). This is in contrast to 14% of XX; dsx^{D}/+ gonads having hubs and 86% having terminal filaments (n=106).
Camta	FBgn0259234	*Camta* is occupied by DSX and is a member of occupancy cluster 1 (Fig 2A). The strength of DSX binding at *Camta* is in the 97th percentile (gene-level PWM score=0.989) of all genes. The conservation of DSX binding at *Camta* is in the 30th percentile (conservation index score=4.365) of all Dmel genes.	Driving a shRNAi against *Camta* (FBst0035683, TRiP GLV21048) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

CG14476	FBgn0027588	*CG14476* is not occupied by DSX and is a member of occupancy cluster 2 (Fig 2A). The strength of DSX binding at *CG14476* is in the 55th percentile (gene-level PWM score=0.207) of all genes. The conservation of DSX binding at *CG14476* could not be determined due to the lack of similar sequences in other Drosophila species.	Driving a shRNAi against *CG14476* (FBst0034334,TRiP HMS01322) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
CG1894	FBgn0039585	*CG1894* is not occupied by DSX and is a member of occupancy cluster 4 (Fig 2A). The strength of DSX binding at *CG1894* is in the 61st percentile (gene-level PWM score=0.251) of all genes. The conservation of DSX binding at *CG1894* is in the 3rd percentile (conservation index score=0.5464) of all Dmel genes.	Driving a shRNAi against *CG1894* (FBst0034925, TRiP HMS01274) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
CG30089	FBgn0050089	*CG30089* is occupied by DSX and is a member of occupancy cluster 2 (Fig 2A). The strength of DSX binding at *CG30089* is in the 95th percentile (gene-level PWM score=0.908) of all genes. The conservation of DSX binding at *CG30089* is in the 88th percentile (conservation index score=30.07) of all Dmel genes.	Driving a shRNAi against *CG30089* (FBst0028321,TRiP JF02953) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

CG30382	FBgn0050382	*CG30382* is not occupied by DSX and is a member of occupancy cluster 1 (Fig 2A). The strength of DSX binding at *CG30382* is in the 43rd percentile (gene-level PWM score=0.146) of all genes. The conservation of DSX binding at *CG30382* could not be determined due to the lack of similar sequences in other Drosophila species.	Driving a shRNAi against *CG30382* (FBst0027557, TRiP JF02711) with *dsx*-Gal4 (FBal0277019), results in late-pupal lethality. No such lethality was observed in *dsx*-Gal4 alone or in Oregon R controls.

chm	FBgn0028387	*chm* is occupied by DSX and is a member of occupancy cluster 3 (Fig 2A). The strength of DSX binding at *chm* is in the 93rd percentile (gene-level PWM score=0.853) of all genes. The conservation of DSX binding at *chm* is in the 62nd percentile (conservation index score=9.38) of all Dmel genes.	Wildtype *chm* promotes genital tract and sex comb development in males as well as gonad and genitalia development in both males and females. Driving a shRNAi against *chm* (FBst0032484 TRiP HMS00487) with *dsx*-Gal4 (FBal0277019) results in adult female gonad and genitalia phenotypes as well as adult male gonad, genitalia, genital tract, and sex comb phenotypes. Adult female *chm* knockdown gonads have germline tumors with >16 nuclei as well as detachment of the terminal filament cells from the germarium. The external genitalia of these knockdown females are amorphous and completely lack identifiable female structures such as the vaginal plate and vaginal teeth. Adult male *chm* knockdown gonads have a bulbous apical testis tip and are not connected to the genital tract as the genital tract including the vas deferens is missing along with the ejaculatory bulb and accessory glands. These knockdown adult males also have fewer sex combs, and the sex combs that are present are pointed and thinner than control flies. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *chm* genetically interacts with *dsx* to promote female-like gonad development. 85% of XX; chm^{14}/+; dsx^{D}/+ gonads have hubs while 15% have terminal filaments (n=47). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).
cindr	FBgn0027598	*cindr* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *cindr* is in the 94th percentile (gene-level PWM score=0.876) of all genes. The conservation of DSX binding at *cindr* is in the 60th percentile (conservation index score=8.975) of all Dmel genes.	Driving a shRNAi against *cindr* (FBst0035670, TRiP GLV21035) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

Cyp4g1	FBgn0010019	*Cyp4g1* is not occupied by DSX and is a member of occupancy cluster 2 (Fig 2A). The strength of DSX binding at *Cyp4g1* is in the 38th percentile (gene-level PWM score=0.125) of all genes. The conservation of DSX binding at *Cyp4g1* could not be determined due to the lack of similar sequences in other Drosophila species.	Driving a shRNAi against *Cyp4g1* (FBst0036737, HMS01628) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
dom	FBgn0020306	*dom* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *dom* is in the 95th percentile (gene-level PWM score=0.904) of all genes. The conservation of DSX binding at *dom* is in the 18th percentile (conservation index score=2.888) of all Dmel genes.	Driving a shRNAi against *dom* (FBst0041674, TRiP HMS02208) with *dsx*-Gal4 (FBal0277019), results in late-pupal lethality. No such lethality was observed in *dsx*-Gal4 alone or in Oregon R controls.

dpp	FBgn0000490	*dpp* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *dpp* is in the 94th percentile (gene-level PWM score=0.866) of all genes. The conservation of DSX binding at *dpp* is in the 95th percentile (conservation index score=54.61) of all Dmel genes.	Wildtype *dpp* promotes gonad and genitalia development in males. Driving a shRNAi against *dpp* (FBst0036779, TRiP JF02455) with *dsx*-Gal4 (FBal0277019) results in male-specific adult genitalia and gonad phenotypes. Female knockdown flies are unaffected. Adult male *dpp* knockdown flies have no external genitalia. The gonads of these knockdown flies are atrophic and have few germline and somatic cells. There is no apparent germline stem cell niche structure as evidence by a lack of a N-cadherin labeled hub. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *dpp* genetically interacts with dsx to promote female-like gonad development. 31% of XX; dpp^{d6}/+; dsx^{D}/+ gonads have hubs while 69% have terminal filaments (n=36). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).

ds	FBgn0000497	*ds* is occupied by DSX and is a member of occupancy cluster 3 (Fig 2A). The strength of DSX binding at *ds* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *ds* is in the 99th percentile (conservation index score=161.1) of all Dmel genes.	Wildtype *ds* promotes gonad and sex comb development in males and genitalia development in both males and females. Driving a shRNAi against *ds* (FBst0032964, TRiP HMS00759) with *dsx*-Gal4 (FBal0277019) results in male-specific gonad and sex comb phenotypes as well as male and female genitalia phenotypes. The gonads of adult male *ds* knockdown flies have a bulbous anterior testis tip, and the majority of the testis is occupied by spermatocytes stages. There appears to be no stage of germ cell development after the spermatocyte stage present. Sex combs from the adult male *ds* knockdown flies have more bristles than controls. The genitalia of these knockdown males are missing the most dorsal portions of the genital arch. Adult female *ds* knockdown genitalia do not open to the exterior of the fly and are missing vaginal teeth primarily in the more ventral portions of the plate. These female genitalia are also not bilaterally symmetric as in normal genitalia. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *ds* does not genetically interact with dsx. 5% of ds^{33K}/+; dsx^{D}/+ gonads have hubs while 95% have terminal filaments (n=38). This is in comparison to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).

dsx	FBgn0000504	*dsx* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *dsx* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *dsx* is in the 98th percentile (conservation index score=95.07) of all Dmel genes.	Wildtype *dsx* promotes gonad, genitalia, genital tract, sex comb, and abdominal pigmentation in males and females. Driving RNAi against *dsx* (FBst0450212, VDRC v110306) with *dsx*-Gal4 (FBal0277019) results in adult female and male gonad, genitalia, genital tract, sex comb, and abdominal pigmentation phenotypes. Adult female *dsx* knockdown genitalia have reduced numbers of vaginal teeth and vaginal plates that are not bilaterally symmetric. Occasionally one or both sides of the vaginal plate is completely missing. These *dsx* knockdown adult females also have gonads that resemble 3rd instar ovaries being rounded and smaller than adult ovaries. These gonads are not connected to the genital tract as the genital tract including the oviduct was missing along with the parovaria and spermathecae. *dsx* knockdown adult females also have intersexual abdominal pigmentation with all of abdominal tergite A6 pigmented and approximately half of A5 pigmented. *dsx* knockdown adult females have intersexual sex combs with ~5 bristles present that resemble male sex combs being pigmented and rotated approximately 45 degrees relative to the transverse rows of bristles. Adult male *dsx* knockdown genitalia are missing several male genital structures including the genital arch, lateral lobes, penis apparatus, and clasper teeth. These *dsx* knockdown adult males have gonads that resemble 3rd instar testes being rounded and not elongated. These gonads are not connected to the genital tract as the genital tract including the vas deferens is completely missing as is the accessory gland and ejaculatory bulb. *dsx* knockdown adult males also have intersexual abdominal pigmentation with all of abdominal tergite A6 pigmented and approximately half of A5 pigmented. *dsx* knockdown adult males have intersexual sex combs with ~5 bristles present that partially resemble male sex combs being pigmented and rotated approximately 45 degrees relative to the transverse rows of bristles. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls.

ear	FBgn0026441	*ear* is not occupied by DSX and is a member of occupancy cluster 1 (Fig 2A). The strength of DSX binding at *ear* is in the 54th percentile (gene-level PWM score=0.199) of all genes. The conservation of DSX binding at *ear* is in the 46th percentile (conservation index score=6.853) of all Dmel genes.	Wildtype *ear* promotes male sex comb development. Driving a shRNAi against *ear* (FBst0034798, TRiP HMS00107) with *dsx*-Gal4 (FBal0277019) results in a sex-specific adult male sex comb phenotype. These sex combs exhibit reduced pigmentation compared to controls, but their morphology and number were unchanged. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls.
Egfr	FBgn0003731	*Egfr* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Egfr* is in the 95th percentile (gene-level PWM score=0.885) of all genes. The conservation of DSX binding at *Egfr* could not be determined due to the lack of similar sequences in other Drosophila species.	Wildtype *Egfr* promotes male gonad and sex comb development. Driving a shRNAi against *Egfr* (FBst0036773, TRiP JF02384) with *dsx*-Gal4 (FBal0277019) results in male-specific adult gonad and sex comb phenotypes. The gonads of these knockdown males have a bulbous anterior testis tip. The sex combs of these *Egfr* knockdown males are thinner than controls and are not organized into a single row. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *Egfr* genetically interacts with *dsx* to promote female-like gonad development. 100% of XX; $Egfr^{f2}$/+; dsx^{D}/+ gonads have hubs while 0% have terminal filaments (n=56). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).

Eip74EF	FBgn0000567	*Eip74EF* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Eip74EF* is in the 97th percentile (gene-level PWM score=0.983) of all genes. The conservation of DSX binding at *Eip74EF* is in the 94th percentile (conservation index score=49.68) of all Dmel genes.	Driving a shRNAi against *Eip74EF* (FBst0029353, TRiP JF02515) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
Eip75B	FBgn0000568	*Eip75B* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Eip75B* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *Eip75B* is in the 61st percentile (conservation index score=9.21) of all Dmel genes.	Driving a shRNAi against *Eip75B* (FBst0035780, TRiP HMS01530) with *dsx*-Gal4 (FBal0277019) results in late-pupal lethality. No such lethality was observed in *dsx*-Gal4 alone or in Oregon R controls.
fal	FBgn0028380	*fal* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *fal* is in the 80th percentile (gene-level PWM score=0.470) of all genes. The conservation of DSX binding at *fal* is in the 86th percentile (conservation index score=26.15) of all Dmel genes.	Driving a shRNAi against *fal* (FBst0035708, TRiP GLV21073) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

fz2	FBgn0016797	*fz2* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *fz2* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *fz2* is in the 99th percentile (conservation index score=142.3) of all Dmel genes.	Wildtype *fz2* promotes gonad development in females and males. Driving a shRNAi against *fz2* (FBst0027568, TRiP JF02722) with *dsx*-Gal4 (FBal0277019) results in adult female and male gonad phenotypes. The adult female knockdown gonads are disorganized with no clear linear progression from early to late stage egg development as in control. The terminal filaments of these knockdown gonads are also disorganized and show no clear grouping near one another as is seen in controls. The terminal filament cell stacks that are present have fewer cells than control. The adult male knockdown gonads exhibit germline stem cell niches (i.e. hub, cyst stem cells, and germline stem cells) that lose apical testis tip localization while still maintaining normal morphology. No similar phenotype was observed in *dsx*-Gal4 alone or in Oregon R controls. *fz2* genetically interacts with *dsx* to promote male-like gonad development. 3% of XX; fz2^{MI02902}/dsx^{D} gonads have hubs while 97% have terminal filaments (n=36). This is in contrast to the phenotype of XX; *dsx*^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).

gpp	FBgn0264495	*gpp* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *gpp* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *gpp* is in the 96th percentile (conservation index score=63.18) of all Dmel genes.	Wildtype *gpp* promotes male sex comb development as well as genitalia, gonad, and genital tract development. Driving a shRNAi against *gpp* (FBst0034842, TRiP HMS00160) with *dsx*-Gal4 (FBal0277019) results in adult female and male genitalia, gonad, and genital tract phenotypes as well as a male-specific sex comb phenotype. The female genitalia of these adult *gpp* knockdown flies do not open completely to the outside of the fly. The vaginal plate is partially recessed into the abdomen of the fly and it missing the majority of vaginal teeth although 2-5 are present. The gonads of these knockdown adult females are small and are not organized in a progressive series of egg stage development as controls. There are egg chambers with greater than 16 nuclei. The germline stem cell niche is compromised as evidenced by a collapse of terminal filaments into a cluster of N-cad positive cells. These gonads are not connected to the genital tract as the genital tract including the oviduct was missing along with the parovaria and spermathecae. The adult male *gpp* knockdown genitalia are missing the majority of the dorsal male genital arch. The gonads of these adult male knockdown flies are rounded resembling 3rd instar testes. These gonads are not connected to the genital tract as the genital tract including the vas deferens is missing in the large majority of flies. The ejaculatory bulb is also missing; however, the majority of flies still have an accessory gland present. There are fewer sex comb bristles in these adult male knockdown flies, and those sex comb bristles that do exist have nearly no pigment; however, all bristles on this tarsal segment are also lacking pigment. Adult female knockdown legs also lack pigment on the first tarsal segment although no sex comb bristles are present. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *gpp* genetically interacts with *dsx* to promote female-like gonad development. 62% of XX; gpp^{X}/dsx^{D} gonads have hubs while 38% have terminal filaments (n=37). This is in

Gug	FBgn0010825	*Gug* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Gug* is in the 85th percentile (gene-level PWM score=0.584) of all genes. The conservation of DSX binding at *Gug* is in the 77th percentile (conservation index score=16.62) of all Dmel genes.	Driving a shRNAi against Gug (FBst0032961, TRiP HMS00756) with dsx-Gal4 (FBal0277019), results in late-pupal lethality. No such lethality was observed in dsx-Gal4 alone or in Oregon R controls. Gug also genetically interacts with *dsx* to promote female-like gonad development. 47% of XX; Gug^{J5A3}/+; dsx^{D}/+ gonads have hubs while 53% have terminal filaments (n=34). This is in contrast to 14% of XX; dsx^{D}/+ gonads having hubs and 86% having terminal filaments (n=106).

H	FBgn0001169	*H* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *H* is in the 40th percentile (gene-level PWM score=0.130) of all genes. The conservation of DSX binding at *H* is in the 75th percentile (conservation index score=15.42) of all Dmel genes.	Wildtype *H* promotes male sex comb development as well as genitalia and gonad development in males and females. Driving a shRNAi against *H* (FBst0034703, TRiP HMS01182) with *dsx*-Gal4 (FBal0277019) results in a male-specific sex comb phenotype as well as male and female genitalia and gonad phenotypes. Adult male *H* knockdown flies have no bristles resembling sex combs on the forelegs. These *H* adult male knockdown flies also have bulbous apical testis tip where the germline stem cell niche (i.e. hub cells, cyst stem cells, and germline stem cells) along with some early gonial cells appear "pinched off" from the rest of the testis. The genitalia of these male *H* knockdown adults are missing the more ventral aspects of the male genitalia including the lateral lobes, clasper teeth, and penis apparatus. The genital arch and anal plate are present but with a rotation defect. Adult female *H* knockdown genitalia are missing all identifiable female genital components missing save an opening that mildly resembles a vaginal plate. These knockdown females also have gonads contain very few late stage egg chambers. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *H* genetically interacts with *dsx* to promote male-like gonad development. 0% of H^{1}/+ dsx^{D}/+ gonads have hubs while 100% have terminal filaments (n=35). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).

hang	FBgn0026575	*hang* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *hang* is in the 92nd percentile (gene-level PWM score=0.794) of all genes. The conservation of DSX binding at *hang* is in the 27th percentile (conservation index score=4.016) of all Dmel genes.	Wildtype *hang* promotes genitalia development in females and genitalia and vas deferens development in males. Driving a shRNAi against *hang* (FBst0035674; TRiP GLV21039 with *dsx*-Gal4 (FBal0277019) results in a sex-specific adult female genitalia phenotype and sex-specific adult male gonad and vas deferens phenotypes. The adult female *hang* knockdown genitalia are recessed into the abdomen and are truncated in the ventral portion of the plate. The germline stem cell niche (i.e. hub cells, cyst stem cells, and germline stem cells) are "pinched off" and protruding from the apical testis tip of adult male *hang* knockdown gonads. The adult male vas deferens is not as wide as *dsx*-Gal4 alone and Oregon R controls. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls.

Hsc70-4	FBgn0001219	*Hsc70-4* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Hsc70-4* is in the 54th percentile (gene-level PWM score=0.205) of all genes. The conservation of DSX binding at *Hsc70-4* is in the 65th percentile (conservation index score=9.492) of all Dmel genes.	Wildtype *Hsc70-4* promotes male sex comb development as well as gonad, genitalia, and genital tract development in males and females. Driving a shRNAi against *Hsc70-4* (FBst0035684; TRiP GLV21049) with *dsx*-Gal4 (FBal0277019) results in a sex-specific adult male sex comb phenotype as well as adult female and male gonad, genitalia, and genital tract phenotypes. Adult female *Hsc70-4* knockdown gonads resemble 3rd instar ovaries being rounded and not organized into individual ovarioles. These gonads are not connected to the genital tract as the genital tract including the oviduct was missing along with the parovaria and spermathecae. The adult female *Hsc70-4* knockdown flies completely lack vaginal plates although the anal plate is present. Adult male *Hsc70-4* knockdown gonads are resemble 3rd instar testes being rounded and not elongated. These knockdown gonads are not connected to the genital tract as the genital tract including the vas deferens is missing along with the ejaculatory bulb and accessory glands. The adult male *Hsc70-4* knockdown flies have external genitalia that are missing the penis apparatus. The tips of the sex combs from these adult knockdown male flies are pointed, rather than rounded, and thinner than controls. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls.

InR	FBgn0013984	*InR* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *InR* is in the 98th percentile (gene-level PWM score=0.995) of all genes. The conservation of DSX binding at *InR* is in the 18th percentile (conservation index score=2.807) of all Dmel genes.	Wildtype *InR* promotes gonad development in males. Driving a shRNAi against *InR* (FBst0031594, TRiP JF01183) with *dsx*-Gal4 (FBal0277019) results in a male-specific adult gonad phenotype. The gonads of these adult male *InR* knockdown flies have a bulbous anterior tip and the germline stem cell niche (i.e. hub cells, cyst stem cells, and germline stem cells) are often "pinched off" and protruding from the apical testis. There is also apparent loss of later germ cell lineages most notably at the spermatocyte stage. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *InR* genetically interacts with *dsx* to promote female-like gonad formation. 29% of InR^{E19}/dsx^{D} gonads have hubs while 71% have terminal filaments (n=35). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).
ix	FBgn0001276	*ix* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *ix* is in the 71st percentile (gene-level PWM score=0.342) of all genes. The conservation of DSX binding at *ix* could not be determined due to the lack of similar sequences in other Drosophila species.	Wildtype *ix* promotes female genitalia development. Driving a shRNAi against *ix* (FBst0037545, TRiP GLV21092) with *dsx*-Gal4 (FBal0277019) results in a female-specific adult genitalia phenotype. The genitalia of these knockdown females are recessed into the abdomen and are not completely formed as they are truncated in the ventral portions of the plate. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls.

jbug	FBgn0028371	*jbug* is occupied by DSX and is a member of occupancy cluster 2 (Fig 2A). The strength of DSX binding at *jbug* is in the 92nd percentile (gene-level PWM score=0.791) of all genes. The conservation of DSX binding at *jbug* is in the 65th percentile (conservation index score=9.492) of all Dmel genes.	Wildtype *jbug* promotes normal female abdominal pigmentation. Driving a shRNAi against *jbug* (FBst0039070, TRiP HMS01990) with *dsx*-Gal4 (FBal0277019), results in loss of posteriolateral A6 tergite pigmentation and all A7 tergite pigmentation in adult female files with males being unaffected. No similar phenotype was observed in *dsx*-Gal4 alone or in Oregon R controls.
Jheh2	FBgn0034405	*Jheh2* is not occupied by DSX and is a member of occupancy cluster 3 (Fig 2A). The strength of DSX binding at *Jheh2* is in the 48th percentile (gene-level PWM score=0.169) of all genes. The conservation of DSX binding at *Jheh2* is in the 34th percentile (conservation index score=4.973) of all Dmel genes.	Driving a shRNAi against *Jheh2* (FBst0035655, TRiP GLV21020) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
jub	FBgn0030530	*jub* is not occupied by DSX and is a member of occupancy cluster 3 (Fig 2A). The strength of DSX binding at *jub* is in the 79th percentile (gene-level PWM score=0.457) of all genes. The conservation of DSX binding at *jub* is in the 12th percentile (conservation index score=1.979) of all Dmel genes.	Driving a shRNAi against *jub* (FBst0041938, TRiP HMS02335) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

lea	FBgn0002543	*lea* is not occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *lea* is in the 95th percentile (gene-level PWM score=0.928) of all genes. The conservation of DSX binding at *lea* is in the 75th percentile (conservation index score=15.75) of all Dmel genes.	Driving a shRNAi against *lea* (FBst0034589, TRiP HMS01063) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
lid	FBgn0031759	*lid* is not occupied by DSX and is a member of occupancy cluster 3 (Fig 2A). The strength of DSX binding at *lid* is in the 90th percentile (gene-level PWM score=0.747) of all genes. The conservation of DSX binding at *lid* is in the 56th percentile (conservation index score=8.354) of all Dmel genes.	Driving a shRNAi against *lid* (FBst0028944, TRiP HM05155) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
ligatin	FBgn0041588	*ligatin* is not occupied by DSX and is a member of occupancy cluster 2 (Fig 2A). The strength of DSX binding at *ligatin* is in the 42th percentile (gene-level PWM score=0.141) of all genes. The conservation of DSX binding at *ligatin* could not be determined due to the lack of similar sequences in other Drosophila species.	Driving a shRNAi against *ligatin* (FBst0033995, TRiP HMS00958) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

lilli	FBgn0041111	*lilli* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *lilli* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *lilli* is in the 94th percentile (conservation index score=50.51) of all Dmel genes.	Wildtype *lilli* promotes gonad and sex comb development in males as well as genitalia development in both sexes. Driving a shRNAi against *lilli* (FBst0034592, TRiP HMS01066) with *dsx*-Gal4 (FBal0277019) results in male-specific adult gonad and sex comb phenotypes as well as male and female adult genitalia phenotypes. The testes of these adult *lilli* knockdown flies are wider than controls likely due to the large number of germ cells within the testis. These knockdown males also have sex combs that have less pigmentation than controls. The genitalia of the adult male *lilli* knockdown flies have a severe rotation defect. All male structures are still present but reduced in size. The adult female *lilli* knockdown flies have genitalia that are missing vaginal teeth in the more ventral portions of the vaginal plate. The vaginal plate is also truncated in the more ventral portions. These genitalia do not open to the outside of the fly. No similar phenotypes were observed in dsx-Gal4 alone or in Oregon R controls. *lilli* genetically interacts with *dsx* to promote female-like gonad development. 39% of XX; $lilli^{A17\text{-}2}$/+; dsx^{D}/+ gonads have hubs while 61% have terminal filaments (n=36). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).

lola	FBgn0005630	*lola* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *lola* is in the 98th percentile (gene-level PWM score=0.997) of all genes. The conservation of DSX binding at *lola* is in the 99th percentile (conservation index score=127.3) of all Dmel genes.	Wildtype *lola* promotes female genitalia development and male gonad development. Driving a shRNAi against *lola* (FBst0035721, TRiP GLV21086) with *dsx*-Gal4 (FBal0277019) results in a female-specific adult genitalia phenotype and a male-specific adult gonad phenotype. Adult *lola* knockdown females have genitalia that are not closed completely and interior portions of the vaginal opening are everted to the outside of the fly. Vaginal plates and vaginal teeth are present. *lola* adult male knockdown flies have a bulbous apical testis tip and aren't as elongated as a wildtype testis. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *lola* does not genetically interact with dsx. 5% of $lola^{e76}$/+ dsx^{D}/+ gonads have hubs while 95% have terminal filaments (n=43). This is in comparison to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).
Mad	FBgn0011648	*Mad* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Mad* is in the 92nd percentile (gene-level PWM score=0.813) of all genes. The conservation of DSX binding at *Mad* is in the 41st percentile (conservation index score=5.978) of all Dmel genes.	Driving a shRNAi against *mad* (FBst0035648, TRiP GLV21013) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes. *Mad* does not genetically interact with dsx. 8% of P{w^{+MC}=lacW} Mad^{k00237}/+; dsx^{D}/+ gonads have hubs while 92% have terminal filaments (n=36). This is in comparison to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).

mam	FBgn0002643	*mam* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *mam* is in the 98th percentile (gene-level PWM score=0.998) of all genes. The conservation of DSX binding at *mam* is in the 92nd percentile (conservation index score=41.29) of all Dmel genes.	Wildtype *mam* promotes male sex comb and abdominal pigmentation development as well as genitalia development in both sexes. Driving a shRNAi against *mam* (FBst0028046, TRiP JF02881) with *dsx*-Gal4 (FBal0277019) results in male-specific sex comb and abdominal pigmentation phenotypes as well as male and female genitalia phenotypes. The sex combs of adult male mam knockdown flies have no bristles resembling sex combs on the forelegs. These knockdown males also have an abdominal pigmentation defect specifically in the A5 tergite such that this tergite is only partially pigmented rather than fully pigmented in controls. The A6 tergite in unaffected. Adult male knockdown flies also have no external genitalia. Adult female *mam* knockdown also lack external genitalia. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *mam* genetically interacts with *dsx* to promote female-like gonad development. 44% of mam^{8}/+ dsx^{D}/+ gonads have hubs while 56% have terminal filaments (n=36). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).
mamo	FBgn0264981	*mamo* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *mamo* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *mamo* is in the 99th percentile (conservation index score=120.5) of all Dmel genes.	Wildtype *mamo* promotes female gonad development and male sex comb development. Driving a shRNAi against *mamo* (FBst0037486, TRiP GLV21091) with *dsx*-Gal4 (FBal0277019), results in a female-specific gonad phenotype and a male-specific sex comb phenotype. Adult female *mamo* knockdown ovaries occupy the majority of the abdomen with the majority of space taken by late stage eggs. Females are capable of laying fewer eggs that controls. Adult male *mamo* knockdown sex combs have teeth that are pointed and thinner than controls. Further, the first 4 proximal bristles of these knockdown sex combs are not aligned in a row with the remaining distal bristles. Total bristle number is not affected by this knockdown. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls.

mgl	FBgn0261260	*mgl* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *mgl* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *mgl* is in the 67th percentile (conservation index score=10.79) of all Dmel genes.	Wildtype *mgl* promotes female genitalia development and male gonad development. Driving a shRNAi against *mgl* (FBst0029324, TRiP JF02485) with *dsx*-Gal4 (FBal0277019) results in an adult female-specific genitalia phenotype and an adult male-specific gonad phenotype. The genitalia of adult female *mgl* knockdown flies are missing the majority of vaginal teeth. The gonads of adult male *mgl* knockdown flies are rounded and not elongated like 3rd instar testes, but these gonads are connected to the genital tract. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *mgl* genetically interacts with *dsx* to promote female-like gonad development. 63% of XX; $P\{EP\}mgl^{G17430}/dsx^{D}$ gonads have hubs while 37% have terminal filaments (n=38). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).).
neur	FBgn0002932	*neur* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *neur* is in the 67th percentile (gene-level PWM score=0.297) of all genes. The conservation of DSX binding at *neur* could not be determined due to the lack of similar sequences in other Drosophila species.	Wildtype *neur* promotes normal sex comb development in males and females. Driving a shRNAi against *neur* (FBst0026023, TRiP JF02048) with *dsx*-Gal4 (FBal0277019) results in adult female and male sex comb phenotypes. Adult female knockdown flies have an intersexual sex comb that consists of the central bristle normally only found in male sex combs. Male *neur* knockdown sex combs are duplicated with more than one row of sex combs present. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *neur* genetically interacts with *dsx* to promote male-like gonad development. 0% of $neur^{11}$/+ dsx^{D}/+ gonads have hubs while 100% have terminal filaments (n=36). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).

Nurf-38	FBgn0016687	*Nurf-38* is not occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Nurf-38* is in the 14th percentile (gene-level PWM score=0.046) of all genes. The conservation of DSX binding at *Nurf-38* could not be determined due to the lack of similar sequences in other Drosophila species.	Wildtype *Nurf-38* promotes gonad, genitalia, and genital tract development in both sexes. Driving a shRNAi against *Nurf-38* (FBst0031341, TRiP JF01299) with *dsx*-Gal4 (FBal0277019) results in adult female and male gonad, genitalia, and genital tract phenotypes. Adult female *Nurf-38* knockdown gonads resemble 3rd instar ovaries being rounded and not organized into individual ovarioles. These gonads are not connected to the genital tract as the genital tract including the oviduct is missing along with spermathecae and parovaria. The adult female *Nurf-38* knockdown flies have no external genitalia. Adult male *Nurf-38* knockdown gonads are small and resemble 3rd instar testes being rounded and not elongated. These gonads are not connected to the genital tract as the genital tract including the vas deferens is missing along with the accessory gland and ejaculatory blub. The adult genitalia of these *Nurf-38* knockdown males exhibit a rotation defect. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls.
Pdcd4	FBgn0030520	*Pdcd4* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Pdcd4* is in the 93rd percentile (gene-level PWM score=0.844) of all genes. The conservation of DSX binding at *Pdcd4* is in the 83rd percentile (conservation index score=22.43) of all Dmel genes.	Driving a shRNAi against *Pdcd4* (FBst0035712, TRiP GLV21077) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

Pde11	FBgn0085370	*Pde11* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Pde11* is in the 96th percentile (gene-level PWM score=0.961) of all genes. The conservation of DSX binding at *Pde11* is in the 98th percentile (conservation index score=80.77) of all Dmel genes.	Driving a shRNAi against *Pde11* (FBst0034611, TRiP HMS01286) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
Plod	FBgn0036147	*Plod* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Plod* is in the 95th percentile (gene-level PWM score=0.924) of all genes. The conservation of DSX binding at *Plod* is in the 19th percentile (conservation index score=2.963) of all Dmel genes.	Driving a shRNAi against *Plod* (FBst0034911, TRiP HMS01259) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

pnt	FBgn0003118	*pnt* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *pnt* is in the 98th percentile (gene-level PWM score=0.993) of all genes. The conservation of DSX binding at *pnt* is in the 78th percentile (conservation index score=17.42) of all Dmel genes.	Wildtype *pnt* promotes male gonad, genital tract, and sex comb development. Driving a shRNAi against *pnt* (FBst0031936, TRiP JF02227) with *dsx*-Gal4 (FBal0277019) results in male-specific gonad, genital tract, and sex comb phenotypes. Adult *pnt* knockdown gonads have a bulbous apical testis tip and the germline stem cell niche (i.e. hub cells, cyst stem cells, and germline stem cells) are often "pinched off" and protruding from the apical testis. The adult male knockdown genital tract is missing the accessory glands and the vas deferens, but the ejaculatory bulb is still present. The adult knockdown male sex combs are improperly rotated relative to the other transverse rows of bristles, yet these sex combs were thicker and shorter than control sex combs. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *pnt* genetically interacts with *dsx* to promote male-like gonad development. 0% of XX; $pnt^{\Delta 88}/dsx^{D}$ gonads have hubs while 100% have terminal filaments (n=43). This is in contrast to 14% of XX; $dsx^{D}/+$ gonads having hubs and 86% having terminal filaments (n=106).
Prosap	FBgn0040752	*Prosap* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Prosap* is in the 98th percentile (gene-level PWM score=0.997) of all genes. The conservation of DSX binding at *Prosap* is in the 94th percentile (conservation index score=46.24) of all Dmel genes.	Wildtype *Prosap* promotes male genitalia development. Driving a shRNAi against *Prosap* (FBst0027284, TRiP JF02596) with *dsx*-Gal4 (FBal0277019), results in a male-specific genitalia defect with a loss of the dorsal and portions of the lateral male genital arch and male lateral lobe. No similar phenotype was observed in *dsx*-Gal4 alone or in Oregon R controls. Adult female *Prosap* knockdown flies are unaffected.

Psn	FBgn0019947	*Psn* is not occupied by DSX and is a member of occupancy cluster 2 (Fig 2A). The strength of DSX binding at *Psn* is in the 80th percentile (gene-level PWM score=0.464) of all genes. The conservation of DSX binding at *Psn* is in the 30th percentile (conservation index score=4.471) of all Dmel genes.	Driving a shRNAi against *Psn* (FBst0038374, TRiP HMS01843) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
puc	FBgn0243512	*puc* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *puc* is in the 90th percentile (gene-level PWM score=0.736) of all genes. The conservation of DSX binding at *puc* is in the 35th percentile (conservation index score=5.07) of all Dmel genes.	Wildtype *puc* promotes male gonad and sex comb development. Driving a shRNAi against *puc* (FBst0034392; TRiP HMS01386) with *dsx*-Gal4 (FBal0277019) results in sex-specific adult male gonad and sex comb phenotypes. The gonads of these knockdown males have a bulbous apical testis tip and the germline stem cell niche (i.e. hub cells, cyst stem cells, and germline stem cells) are often "pinched off" and protruding from the apical testis. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls.
pyd3	FBgn0037513	*pyd3* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *pyd3* is in the 93rd percentile (gene-level PWM score=0.856) of all genes. The conservation of DSX binding at *pyd3* is in the 82nd percentile (conservation index score=21.14) of all Dmel genes.	Driving a shRNAi against *pyd3* (FBst0034557, TRiP HMS01029) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

RapGAP1	FBgn0264895	*RapGAP1* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *RapGAP1* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *RapGAP1* is in the 99th percentile (conservation index score=116.1) of all Dmel genes.	Driving a shRNAi against *RapGAP1* (FBst0031250, TRiP JF01766) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
RunxA	FBgn0083981	*RunxA* is not occupied by DSX and is a member of occupancy cluster 4 (Fig 2A). The strength of DSX binding at *RunxA* is in the 95th percentile (gene-level PWM score=0.923) of all genes. The conservation of DSX binding at *RunxA* is in the 97th percentile (conservation index score=75.81) of all Dmel genes.	Driving a shRNAi against *RunxA*, (FBst0033353TRIP HMS00224) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
rut	FBgn0003301	*rut* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *rut* is in the 98th percentile (gene-level PWM score=0.993) of all genes. The conservation of DSX binding at *rut* is in the 98th percentile (conservation index score=88.85) of all Dmel genes.	Driving a shRNAi against *rut* (FBst0027035, TRiP JF02361) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

salm	FBgn0261648	*salm* is not occupied by DSX and is a member of occupancy cluster 3 (Fig 2A). The strength of DSX binding at *salm* is in the 61st percentile (gene-level PWM score=0.250) of all genes. The conservation of DSX binding at *salm* could not be determined due to the lack of similar sequences in other Drosophila species.	Wildtype *salm* promotes gonad and genitalia development in males. Driving a shRNAi against *salm* (FBst0033714, TRiP HMS00594) with *dsx*-Gal4 (FBal0277019), results in male-specific gonad and genitalia phenotypes. Adult *salm* knockdown male gonads showed clear loss of later germline stages, most notably spermatocyte stages. The external genitalia of the adult *salm* knockdown males are missing male structures including the clasper teeth and penis apparatus. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *salm* genetically interacts with *dsx* to promote female-like gonad development. 85% of XX; *P{PZ} salm*/+; dsx^{D}/+ gonads have hubs while 15% have terminal filaments (n=56). 83% of XX; $salm^{1}$/+; dsx^{D}/+ gonads have hubs while 17% have terminal filaments (n=41). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).

sbb	FBgn0010575	*sbb* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *sbb* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *sbb* could not be determined due to the lack of similar sequences in other Drosophila species.	Wildtype *sbb* promotes genital tract development in females; sex comb and abdominal pigmentation development in males; and gonad and genitalia development in both sexes. Driving a shRNAi against *sbb* (FBst0027049, TRiP JF02375) with *dsx*-Gal4 (FBal0277019) results in adult female gonad, genitalia, and genital tract phenotypes as well as adult male gonad, genitalia, abdominal pigmentation, and sex comb phenotypes. Adult female *sbb* knockdown gonads occupy the majority of the abdomen with stage 14 eggs taking the majority of the space. These gonads have defects in their germline stem cell niche structure as evidenced by the collapse of terminal filaments in to a cluster rather than being elongated in a stack as in controls. In some instances, there is loss of N-cad labeled terminal filament cells at the tip of the ovary. These gonads are not connected to the genital tract as the genital tract including the oviduct was missing along with the parovaria and spermathecae. The adult female *sbb* knockdown flies have poorly-formed genitalia that lack vaginal plates and vaginal teeth. Adult male *sbb* knockdown gonads are wider than control male testes. This is especially apparent near the terminal epithelium. These testes are connected to the genital tract, and the tract appears normal. The adult genitalia of these *sbb* knockdown flies have a rotation defect. These genitalia are also missing the penis apparatus. *sbb* knockdown adult males have female-like pigmentation in abdominal tergite A5 with roughly the posterior 40% of the A5 tergite being pigmented. Pigmentation in A6 is unaffected. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls.

Set1	FBgn0040022	*Set1* is occupied by DSX and is a member of occupancy cluster 3 (Fig 2A). The strength of DSX binding at *Set1* is in the 87th percentile (gene-level PWM score=0.645) of all genes. The conservation of DSX binding at *Set1* could not be determined due to the lack of similar sequences in other Drosophila species.	Wildtype *Set1* promotes genitalia development in females and sex comb development in males. Driving a shRNAi against *Set1* (FBst0033704; TRiP HMS00581) with *dsx*-Gal4 (FBal0277019) results in a sex-specific adult female genitalia phenotype and a sex-specific adult male sex comb phenotype. Adult female *Set1* knockdown genitalia are missing the vaginal plate, but the anal plate is still present. Male genitalia are unaffected. The tips of sex combs from adult male *Set1* knockdown flies are pointed, rather than rounded as in controls, and thinner than controls. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls.
skpC	FBgn0026175	*skpC* is not occupied by DSX and is a member of occupancy cluster 4 (Fig 2A). The strength of DSX binding at *skpC* is in the 31st percentile (gene-level PWM score=0.099) of all genes. The conservation of DSX binding at *skpC* is in the 1st percentile (conservation index score=0.1216) of all Dmel genes.	Driving a shRNAi against *skpC* (FBst0033925, TRiP HMS00871) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
sli	FBgn0264089	*sli* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *sli* is in the 97th percentile (gene-level PWM score=0.982) of all genes. The conservation of DSX binding at *sli* is in the 85th percentile (conservation index score=25.64) of all Dmel genes.	Wildtype *sli* promotes gonad and accessory gland development in males. Driving a shRNAi against *sli* (FBst0031467; TRiP JF01228) with *dsx*-Gal4 (FBal0277019) results in adult male-specific gonad and accessory gland phenotypes. The gonads of these adult male *sli* knockdowns have a bulbous and large apical testis tip and are wider than control testes. These knockdown males also have accessory glands that are longer and wider than controls. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls.

Smr	FBgn0265523	*Smr* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Smr* is in the 98th percentile (gene-level PWM score=0.998) of all genes. The conservation of DSX binding at *Smr* is in the 93th percentile (conservation index score=45.96) of all Dmel genes.	Wildtype *Smr* promotes normal abdominal pigmentation in females; sex comb development in males; and genitalia, gonad, and gential tract development in both sexes. Driving a shRNAi against *Smr* (FBst0027068, TRiP JF02413) with *dsx*-Gal4 (FBal0277019) results in a female-specific abdominal pigmentation phenotype, a male-specific sex comb phenotype, and female and male genitalia, gonad, and genital tract phenotypes. In adult female *Smr* knockdown flies, the A5 and A6 abdominal tergites were completely pigmented and resembled the normal male pigmentation pattern. The genitalia of adult female *Smr* knockdown flies are completely missing the vaginal plate, but the anal plate is still present. The gonads of these female knockdown flies had a range of phenotypes with some gonads being occupied primarily with germ cells and no terminal filaments being present. Other knockdown goads were small with egg chambers that had greater than 16 nuclei and other egg chambers that were atrophic. In gonads with stage 14 eggs, the dorsal appendages were malformed. These gonads are not connected to the genital tract as the genital tract including the oviduct was missing along with the parovaria and spermathecae. Adult male *Smr* knockdown sex combs are not aligned into a single row like control sex combs. Knockdown male genitalia have a rotation defect. The more ventral male structures like the lateral lobes, clasper teeth, and penis apparatus are missing while the dorsal male structures such as the genital arch are still present. The anal plate is unaffected. The gonads of these male knockdown flies resemble 3rd instar-like gonads being rounded and not elongated like adult testes. These gonads are not connected to the genital tract because the genital tract including the vas deferens is missing along with the ejaculatory bulb and accessory glands. No similar phenotypes were observed in dsx-Gal4 alone or in Oregon R controls. *Smr* genetically interacts with *dsx* to promote male-like gonad

Snoo	FBgn0085450	*Snoo* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Snoo* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *Snoo* is in the 99th percentile (conservation index score=141.8) of all Dmel genes.	Wildtype *Snoo* promotes male gonad development. Driving a shRNAi against *Snoo* (FBst0031934, TRiPJF02225) with *dsx*-Gal4 (FBal0277019) results in a sex-specific adult male gonad phenotype. These gonads have a bulbous apical testis tip with excessive numbers of germ cells. The germline stem cell niche (i.e. hub cells, cyst stem cells, and germline stem cells) appears normal. These gonads are wider than control male testes. This is especially apparent near the terminal epithelium where testes are much wider than controls. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls.
Socs36E	FBgn0041184	*Socs36E* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Socs36E* is in the 90th percentile (gene-level PWM score=0.722) of all genes. The conservation of DSX binding at *Socs36E* is in the 6th percentile (conservation index score=1.044) of all Dmel genes.	Driving a shRNAi against *Socs36E* (FBst0035036, TRiP HMS01450) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

Su(Tpl)	FBgn0014037	*Su(Tpl)* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *Su(Tpl)* is in the 95th percentile (gene-level PWM score=0.927) of all genes. The conservation of DSX binding at *Su(Tpl)* is in the 65th percentile (conservation index score=10.04) of all Dmel genes.	Wildtype *Su(Tpl)* promotes male gonad and sex comb development as well as genital development in both sexes. Driving a shRNAi against *Su(Tpl)* (FBst0033399, TRiP HMS00277) with *dsx*-Gal4 (FBal0277019) results in male-specific gonad and sex comb phenotypes as well as male and female genitalia phenotypes. Adult male *Su(Tpl)* knockdown flies have a bulbous and rounded apical testis tip. These knockdown males also have fewer sex comb bristles that are also less pigmented than controls. Adult male knockdown genitalia are recessed into the abdominal cavity and are reduced in overall size, but all structural components are present. Adult female *Su(Tpl)* genitalia are not closed completely resulting in interior portions of the vaginal opening being everted to the outside of the fly. Vaginal plates are present but no vaginal teeth are present. No similar phenotypes were observed *in dsx*-Gal4 alone or in Oregon R controls. Su(Tpl) genetically interacts with dsx to promote female-like gonad development. 65% of $Su(Tpl)^{10}/dsx^{D}$ gonads have hubs while 35% have terminal filaments (n=37). This is in contrast to 14% of XX; dsx^{D}/+ gonads having hubs and 86% having terminal filaments (n=106).
tai	FBgn0041092	*tai* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *tai* is in the 99th percentile (gene-level PWM score=0.999) of all genes. The conservation of DSX binding at *tai* is in the 99th percentile (conservation index score=174.9) of all Dmel genes.	Wildtype *tai* promotes male gonad development. Driving a shRNAi against *tai* (FBst0028971, TRiP HM05182) with *dsx*-Gal4 (FBal0277019), results in a male-specific gonad phenotype where the germline stem cell niche (i.e. hub, cyst stem cells, and germline stem cells) loses apical testis tip localization while still maintaining normal morphology. No similar phenotype was observed in *dsx*-Gal4 alone or in Oregon R controls. Adult female knockdown flies are unaffected. *tai* does not genetically interact with *dsx*. 5% of XX; $P\{lacW\}tai^{k15101}/dsx^{D}$ gonads have hubs while 95% have terminal filaments (n=38). This is in comparison to 14% of XX; dsx^{D}/+ gonads having hubs and 86% having terminal filaments (n=106).

term	FBgn0003683	*term* is not occupied by DSX and is a member of occupancy cluster 4 (Fig 2A). The strength of DSX binding at *term* is in the 8th percentile (gene-level PWM score=0.030) of all genes. The conservation of DSX binding at *term* could not be determined due to the lack of similar sequences in other Drosophila species.	Driving a shRNAi against *term* (FBst0035681, TRiP GLV21046) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
tj	FBgn0000964	*tj* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *tj* is in the 80th percentile (gene-level PWM score=0.476) of all genes. The conservation of DSX binding at *tj* is in the 74th percentile (conservation index score=14.91) of all Dmel genes.	Wildtype *tj* promotes female gonad development. Driving a shRNAi against *tj* (FBst0034595, TRiP HMS01069) with *dsx*-Gal4 (FBal0277019) results in a female-specific ovary phenotype. The ovaries of these knockdown flies are extremely small compared to controls. Frequently, no identifiable ovary is found. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *tj* genetically interacts with dsx to promote male-like gonad development. 2% of XX; tj^{PL3}/+; dsx^{D}/+ gonads have hubs while 98% have terminal filaments (n=44). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).

tkv	FBgn0003716	*tkv* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *tkv* is in the 97th percentile (gene-level PWM score=0.988) of all genes. The conservation of DSX binding at *tkv* is in the 32th percentile (conservation index score=4.696) of all Dmel genes.	Wildtype *tkv* promotes normal female abdominal pigmentation. Driving a shRNAi against *tkv* (FBst0040937, TRiP HMS02185) with *dsx*-Gal4 (FBal0277019) results in a female-specific adult abdominal pigmentation phenotype. The A5 and A6 abdominal tergites are fully pigmented in these adult female *tkv* knockdown flies, and are nearly indistinguishable from male controls. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *tkv* genetically interacts with *dsx* to promote female-like gonad development. 54% of XX; tkv^{8}/+; dsx^{D}/+ gonads have hubs while 46% have terminal filaments (n=36). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).
tll	FBgn0003720	*tll* is not occupied by DSX and is a member of occupancy cluster 4 (Fig 2A). The strength of DSX binding at *tll* is in the 43rd percentile (gene-level PWM score=0.143) of all genes. The conservation of DSX binding at *tll* is in the 36th percentile (conservation index score=5.173) of all Dmel genes.	Driving a shRNAi against *tll* (FBst0034329, TRiP HMS01316) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.

wb	FBgn0261563	*wb* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *wb* is in the 98th percentile (gene-level PWM score=0.992) of all genes. The conservation of DSX binding at *wb* is in the 49th percentile (conservation index score=7.259) of all Dmel genes.	Wildtype *wb* promotes female genitalia development. Driving a shRNAi against *wb* (FBst0029559, TRiP JF03238) with *dsx*-Gal4 (FBal0277019) results in female-specific genitalia defect whereby the genitalia are recessed. Further, the two sides of the vaginal plate are not symmetrical. No similar phenotypes were observed in dsx-Gal4 alone or in Oregon R controls. *wb* genetically interacts with *dsx* to promote female-like gonad development. 32% of XX; $wb^{BG02232}/dsx^{D}$ gonads have hubs while 68% have terminal filaments (n=38). This is in contrast to the phenotype of XX; dsx^{D}/+ where 14% of gonads have hubs and 86% have terminal filaments (n=106).
Wbp2	FBgn0036318	*Wbp2* is not occupied by DSX and is a member of occupancy cluster 3 (Fig 2A). The strength of DSX binding at *Wbp2* is in the 77th percentile (gene-level PWM score=0.420) of all genes. The conservation of DSX binding at *Wbp2* could not be determined due to the lack of similar sequences in other Drosophila species.	Driving a shRNAi against *Wbp2* (FBst0034603, TRiP HMS00563) with *dsx*-Gal4 (FBal0277019) results in no genitalia, gonad, sex comb, abdominal pigmentation, or genital tract phenotypes.
wts	FBgn0011739	*wts* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *wts* is in the 91st percentile (gene-level PWM score=0.765) of all genes. The conservation of DSX binding at *wts* is in the 81st percentile (conservation index score=18.87) of all Dmel genes.	Wildtype *wts* promotes male gonad development. Driving a shRNAi against *wts* (FBst0027662, TRiP JF02741) with *dsx*-Gal4 (FBal0277019) results in a male-specific adult gonad phenotype. The gonads of these adult male knockdown flies resemble 3rd instar-like gonads being rounded and not elongated like adult gonads; however, they are connected to genital tract. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *wts* genetically interacts with *dsx* to promote female-like gonad development. 54% of $wts^{3\text{-}17}/dsx^{D}$ gonads have hubs while 46% have terminal filaments (n=37). This is in contrast to 14% of XX; dsx^{D}/+ gonads having hubs and 86% having terminal filaments (n=106).

yki	FBgn0034970	*yki* is not occupied by DSX and is a member of occupancy cluster 3 (Fig 2A). The strength of DSX binding at *yki* is in the 81st percentile (gene-level PWM score=0.495) of all genes. The conservation of DSX binding at *yki* is in the 51st percentile (conservation index score=7.549) of all Dmel genes.	Driving a shRNAi against *yki* (FBst0031965, TRiP JF03119) with *dsx*-Gal4 (FBal0277019), results in late-pupal lethality. No such lethality was observed in *dsx*-Gal4 alone or in Oregon R controls.
zfh1	FBgn0004606	*zfh1* is occupied by DSX and is a member of occupancy cluster 5 (Fig 2A). The strength of DSX binding at *zfh1* is in the 95th percentile (gene-level PWM score=0.910) of all genes. The conservation of DSX binding at *zfh1* is in the 93rd percentile (conservation index score=43.71) of all Dmel genes.	Wildtype *zfh1* promotes male genitalia and gonad development. Driving a shRNAi against *zfh1* (FBst0029347, TRiP JF02509) with *dsx*-Gal4 (FBal0277019) results in male-specific adult genitalia and gonad phenotypes. The gonads of these adult *zfh1* knockdown males exhibit an overproliferation of spermatocytes as well as a bulbous anterior tip of the testis. There is also loss of later germline lineages occurring after the spermatocyte stage. The genitalia of these knockdown males protrude from the abdomen, but all male structures are present. No similar phenotypes were observed in *dsx*-Gal4 alone or in Oregon R controls. *zfh1* genetically interacts with *dsx* to promote male-like gonad development. 0% of $P\{PZ\}zfh1^{00865}/dsx^{D}$ gonads have hubs while 100% have terminal filaments (n=40). This is in contrast to 14% of XX; dsx^{D}/+ gonads having hubs and 86% having terminal filaments (n=106).

Immunohistochemistry

Tissue was dissected from adult flies aged 1 to 3 days in PBS followed by fixation for 12-25 minutes in PBS containing 0.1% Triton X-100 (PBTx) with 4.0-4.5 % formaldehyde. Samples were blocked in PBTx with 0.1 or 1.0% BSA (BBTx) with or without 2% normal goat serum (NGS) for at least 1 hour and then incubated in BBTx with primary antibody 1-2 hours at room temperature or overnight at 4 °C. Following 3X 10 minute washes in PBTx, samples were incubated in BBTx with or without 2% NGS plus secondary antibody for 1-2 hours at room temperature. Following 3X 10 minute washes in PBTx, samples were mounted on slides in 2.5% DABCO (Sigma-Aldrich, St. Louis, MO, USA) or Fluoromount G (Southern Biotech, Birmingham, AL, USA).

The following primary antibodies were used: chicken anti-VASA (K. Howard) at 1:10,000; rabbit anti-VASA (R. Lehmann) at 1:1000; rat anti-DN-cadherin Ex#8 (Developmental Studies Hybridoma Bank, (Iwai et al., 1997)) at 1:20-50; mouse anti-FAS3 7G10 (Developmental Studies Hybridoma Bank, (Patel et al., 1987)) at 1:30; guinea pig anti-TJ at 1:1000 (Jemc et al., 2012); mouse anti-myc9E10 (Roche); mouse anti-V5 (Invitrogen) at 1:200. The following secondary antibodies were used: Alexa 546 goat anti-chicken at 1:500; Alexa 633 goat anti-chicken at 1:500; Alexa 488 goat anti-rat at 1:300-500; Alexa 633 goat anti-mouse at 1:500; Alexa 546 goat anti-mouse at 1:500; Alexa 488 goat anti-mouse at 1:300; Alexa 488 goat anti-guinea pig at 1:500, Alexa 555 donkey anti-rabbit at 1:300 (Invitrogen, Carlsbad, CA, USA), and Cy5 goat anti-guinea pig (Jackson ImmunoResearch, West Grove, PA, USA) at 1:300. We stained DNA with DAPI (Sigma-Aldrich, St. Louis, MO, USA) at 10 μg/ml for 10 minutes and

then rinsed 3 X 5 minutes in PBTx. All immunohistochemistry samples were imaged on a LSM 510 Meta confocal microscope (Zeiss, Jena, Thuringia, Germany).

dsx^D Genetic interaction screen and candidate gene tests

dsx^D, e^1, Sb^1/TM6B males were crossed to virgin females carrying alleles or deficiencies being tested for genetic interaction with *dsx*. Female offspring of this cross carrying both dsx^D and the allele/deficiency being tested were scored under dissecting microscope for novel phenotypes in abdominal pigmentation, sex comb structure, or genitalia when compared to their female siblings heterozygous for the dsx^D allele. For the deficiency screen, 19 of 101 tested deficiencies shifted some aspect of the external sexual morphology relative to XX; dsx^D/+. Overlapping and adjacent deficiencies with similar effects on the dsx^D/+ phenotype were merged into new shifting regions. When 2 deficiencies overlapped but did not show coordinate genetic interactions, it was assumed that the overlapping region did not contain the gene(s) in each deficiency that interacted with dsx^D; therefore, the overlap was removed. Shifting regions overlapping with non-shifting deficiencies were removed to define 17 unique genomic intervals that shift some aspect of the dsx^D/+ phenotype.

For gonad genetic interaction tests with dsx^D, gonads of female offspring carrying both dsx^D and the allele being tested were dissected. The morphology of mutant gonads was scored by the presence of male or female germline stem cell niches when compared to their female siblings carrying only the dsx^D allele. A gonad was scored as positive for a male germline niche (i.e. hub) if a structure with morphology similar to a normal hub also double-labeled with N-cadherin (N-cad) and Fasciclin III (FasIII). A gonad was scored as positive for a female

germline stem cell niche if a structure with morphology similar to terminal filaments was labeled with N-cad but was not labeled with FasIII (FasIII does not label wildtype terminal filaments). At least 35 gonads were scored for each tested allele.

For RNAi tests of genes, *dsx*-GAL4, (Rideout et al., 2010) and/or (Robinett et al., 2010), virgin females were crossed to males carrying various individual RNAi constructs and raised at 29°C or 25°C. 10 male and 10 female day 3-5 progeny carrying both *dsx*-GAL4 and *UAS-RNAi* were screened under dissecting microscope for phenotypes in the following sexually dimorphic structures: genitalia, gonad, sex comb, abdominal pigmentation, male reproductive tract (accessory gland, ejaculatory duct, ejaculatory bulb), and female reproductive tract (oviduct, spermathecae, parovaria). Flies carrying *dsx*-GAL4 and *UAS-RNAi* were compared to control flies carrying *dsx*-GAL4 alone and Oregon R. For a full description of RNAi alleles and phenotypes see Table S4.

Scanning Electron Microscopy

Adult flies aged 2 days were mounted (without desiccation or other treatment) on aluminum pedestals and imaged in a FEI Quanta 200 ESEM at 80Pa.

DamID-seq and DamID-array

The *UAS-Dam-myc-dsx$^{M/F}$* DNA constructs were made by ligating PCR amplified sex-specific *dsx* cDNA (gift from Gyunghee Lee, University of Tennessee, Knoxville, USA) into the pUASt-attB-NDam-myc plasmid (gift from Tony Southall, University of Cambridge, Cambridge, UK). UASt-attB-Dam-myc (Dam), UASt-attB-Dam-myc-dsx^{F} (*Dam-dsx^{F})*, and UASt-attB-Dam-myc-dsx^{M} (Dam-dsx^{M}) constructs were independently integrated (Genetic Services, Inc.,

Cambridge, MA, USA) into the attP2 site on chromosome 3L using ϕC31 site-directed integration (Bischof et al., 2007). In accordance with DamID protocols, a GAL4 driver was not used in order to keep Dam, Dam-*dsx*^{F} and Dam-*dsx*^{M} expression low to prevent lethality and saturation (Greil et al., 2006; Southall and Brand, 2007). S2 cells were transfected with pMT5.1-DSXM-V5-HisB or pMT5.1-DSXF-V5-HisB (Garrett-Engele et al., 2002) and pCoBlast (Invitrogen, Carlsbad, CA, USA) as the selection plasmid using Effectene (Qiagen, Valencia, CA, USA). Expression was induced using Cu^{+} and presence of fusion proteins was confirmed by immunostaining (Figure S1) and western blot (data not shown). Chromatin immunoprecipitation was performed with anti-V5 tag monoclonal antibody (Invitrogen, Carlsbad, CA, USA) on Protein G coupled Dynabeads (Invitrogen, Carlsbad, CA, USA) followed 1% formaldehyde and shearing chromatin to 200-1000 bp. Adult fatbody and ovaries were dissected from 20-70 flies in PBS at room temperature from day 5 adult flies heterozygous for one of Dam insertions. Samples were transferred to ice after 30 minutes.

Genomic DNA was extracted using components of Qiagen's DNEasy Blood and Tissue kit (Qiagen, Valencia, CA, USA). Samples processed for DamID-seq were homogenized in 175 µl PBS and incubated with 200 mg of RNAse A for 2 minutes at room temperature. Tissue was lysed with 20 µl of proteinase K and 200 µl of buffer AL for ten minutes at 70 °C. 200 µl of ethanol were added to each sample and they were transferred to the spin columns after which genomic DNA extraction continued following manufacturer's instructions (Qiagen, Valencia, CA, USA) with the exception of a 30 minute incubation prior to first elution and a second elution step after a 10 minute incubation. Genomic DNA for DamID-array was extracted following manufacturer's protocol except for the following modifications: a 1.5 hr incubation with lysis buffer prior to the addition of proteinase K, addition of 400 µl of Buffer AL and 300 µl of 100%

ethanol, two rounds of both AW1 and AW2 and an incubation with the elution buffer for 30 minutes prior to two rounds of elution. 2.5 - 3 μg of fatbody genomic DNA and 0.3 μg of ovary genomic DNA was used for selective PCR amplification of methylated DNA. DNA was incubated with 10-30 units of *Dpn*I in 50-100 ul of Buffer 4 (New England Biolabs, Ipswich, MA, USA). *Dpn*I (New England Biolabs, Ipswich, MA, USA) was inactivated at 80 °C for 20 minutes and digested DNA was purified through a Qiaquick PCR Purification column (Qiagen, Valencia, CA, USA) following manufacturer's protocol and eluted in 30 μl ddH20. One-half of the *DpnI* reaction products were ligated to 40 pmol of the doublestranded DamID adaptors (top strand : 5'-CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGA-3'; bottom strand: 5'-TCCTCGGCCG-3') for 2 hours at 16 °C with 400 units of T4 ligase (New England Biolabs, Ipswich, MA, USA) or 5 units of T4 ligase (Roche, Indianapolis, IN, USA) in a 20 μl reaction volume. All 20 μl of the adapter-ligated DNA were then subjected to *Dpn*II digestion with 10 units of *Dpn*II (New England Biolabs, Ipswich, MA, USA) in a 80 μl reaction volume for at least one hour. PCR amplification was performed with 20 μl of the *Dpn*II digested DNA in an 80 μl volume with 100 pmol PCR primer (5′-TCCTCGGCCG-3′), 16 nmol of each dNTP and 1.6 μl PCR Advantage enzyme mix in 1X PCR Advantage Reaction Buffer (Clontech, Mountain View, CA, USA) or 62.5 pmol PCR primer (5′-TCCTCGGCCG-3′), 16 nmol of each dNTP, 80 nmol $MgCl_2$ in 1X buffer with 8 units of *taq* polymerase (Fermentas, Pittsburgh, PA, USA). DNA was amplified with the following program: 10 minutes at 68 °C, 1 minute at 94 °C, 5 minutes at 65 °C and 15 minutes at 68 °C, followed by 3 cycles of 1 minute at 94 °C, 1 minute at 65 °C and 10 minutes at 68 °C and then 17 cycles of 1 minute at 94 °C, 1 minute at 65 °C and 2 minutes at 68 °C. DNA was purified through a Qiaquick column (Qiagen, Valencia, CA, USA).

DamID-array samples were analyzed at Nimblegen where Cy3- and Cy5-fluorescently labeled DamID-prepared DNA was hybridized to the DM_5_Catalog_tiling_HX1 whole genome tiling array and fluorescence data was collected by Roche NimbleGen (Madison, WI, USA). Probe sequence, probe position information and array details are available under GEO accession GPL10639. Three independent biological replicates were collected for DamID-array samples. The raw probe intensity data for each DamID-array experiment was accessed using the DM_5_Catalog_tiling_HXI_pair.txt file provided by Nimblegen (Madison, WI, USA). One dye-flip was performed for each sex and tissue. Arrays were quantile normalized with the R package *preprocessCore* (http://www.bioconductor.org/packages/release/bioc/html/preprocessCore.html). In order to define a lower limit for detection of hybridization, we calculated the mean fluorescent intensity of the 15,758 random sequence probes on the array. The 95th percentile of the mean for random probes was used as a cut-off for hybridization detection. When the replicate means of both Dam-DSX treatment and Dam-only control probes were at or below this value, the data from that probe was removed from the analysis. In order to identify probes with significantly different levels of fluorescent intensity, modified two-sided t-tests were performed assuming unequal variance. p-values were adjusted for multiple testing using the FDR method of Benjamini and Hochberg (Benjamini and Hochberg, 1995). Results of the statistical test for all probes is available at GEO (GSE49480). All calculations and statistical tests were performed in R (R Core Team, 2013). Probes that displayed a FDR < 0.01 and log2Fold-change > 0 were selected for further analysis. When the chromosomal positions of the selected probes positions occurred within 1000 bp of one another they were merged into features to form peaks using BedTools v2.16.2 (Quinlan and Hall, 2010).

For DamID-seq samples, PCR-amplified DNA was sonicated in a 200 µl volume of Qiagen's EB buffer in a BioRuptor Sonicator (Diagenode, Denville, NJ, USA) set on high for 3 X 15 minutes in a 4 °C water bath. Following sonication, DNA was purified through a Qiaquick column (Qiagen, Valencia, CA, USA). Two independent biological replicates were collected for DamID-seq samples. 20 ng of sonicated DamID-prepared DNA were used to make libraries following the protocol in the Illumina ChIP-seq Sample Preparation Kit (Illumina, San Diego, CA, USA). A gel slice of 250-350 bp was excised from the gel prior to PCR amplification. Library concentration was measured on a Nanodrop (Thermo Scientific, Waltham, MA, USA) and size distribution was assessed on a Bioanalyzer (Agilent, Santa Clara, CA, USA). DamID-seq samples were sequenced on a GAIIx or HighSeq 2000 instrument with 76 bp read lengths.

DamID-seq reads were generated using the Illumina pipeline 1.6.47.1 (Male fat body Dam-DSX^{M}), 1.8.70.0 (male fat body Dam-alone, female fat body Dam-alone, female fat body Dam-DSX^{F}), or 1.12.4 (ovary Dam-alone and ovary Dam-DSX^{F}). Reads were trimmed by 17 bp on each end to remove primer sequence and mapped to the *D. melanogaster* genome (FlyBase release 5 with no Uextra) using Bowtie 0.12.7 (Langmead et al., 2009) accepting only uniquely mapped reads with no more than 2 mismatches (-m1 –v2). Duplicate reads were removed from the libraries before peak calling with the Picard tool MarkDuplicates v1.95 (http://picard.sourceforge.net). In order to identify regions of the genome enriched for DSX occupancy, the number of reads occurring in non-overlapping consecutive 500 bp intervals across the genome were counted with HTSeq v0.5.1p2 (Anders et al., 2014). DESeq v1.12.0 (Anders and Huber, 2010) was used for library size normalization and identification of bins significantly enriched for Dam-DSX reads compared to Dam-Only reads (method adapted from (Ross-Innes et al., 2012)). The depth-normalized occupancy signal averaged between replicates, fold changes

and associated p-values and Benjamini-Hochberg FDR-adjusted p-values for each bin are available on the GEO record GSE49480. Bins that contained no reads in either control or treatment samples were removed from the analysis. Bins selected for further analysis for all samples were those that displayed differential read counts with a FDR (Benjamini and Hochberg, 1995) < 0.01 and a log2 Fold-Change > 0 where the number of Dam-DSX reads was the numerator and number of Dam-Only Control reads was the denominator. Adjacent selected bins were combined into features to produce peaks using BEDTools v2.16.2 (Quinlan and Hall, 2010). In order to calculate genome-wide DSX DamID signal (used to create gene level occupancy scores; see below), the log2 Fold Change [(DamDSX +1)/(DamOnly +1)] was calculated for all 500 bp bins across the genome.

ChIP-seq

Schneider Drosophila line 2 cells (S2) were maintained at 25°C in Schneider Drosophila medium (Invitrogen, Carlsbad, CA, USA) containing 10% heat-inactivated Fetal Bovine Serum (JRH Biosciences, Lenexa, KS, USA) and antibiotics (0.5 U/ml penicillin and 0.5 μg/ml streptomycin, Invitrogen, Carlsbad, CA, USA). Cells were transfected with 1 μg expression plasmids (pMT5.1-DSXM-V5-His B and pMT5.1-DSXF-V5-His B (Garrett-Engele et al., 2002) using Effectene Transfection Reagent (Qiagen, Valencia, CA, USA) with 50 ng pCoBlast (Invitrogen, Carlsbad, CA, USA) as the selection plasmid. Following transfection, cells were grown in Schneider Drosophila medium for 60 hours prior to selection with 30 μg/ml blasticidin (Invitrogen, Carlsbad, CA, USA). After 5 weeks of selection, blasticidin-resistant cells were maintained in complete Schneider cell medium containing 25 μg/ml blasticidin. Expression of the recombinant proteins from the MT promoter was induced by adding copper sulfate to the

medium to a final concentration of 500 μM. Presence of the DSX fusion proteins was confirmed by immunostaining (Figure S1) and western blot (data not shown).

~2.7 X 10^8 cells were fixed in 1% formaldehyde for 10 minutes at room temperature. The reaction was quenched by adding glycine to a final concentration of 125 mM and a 5-minute incubation on a shaker at room temperature. Subsequently, the cells were washed twice with ice-cold PBS. After centrifugation at 500 x g (1680 rpm) for 5 min at 4 °C the cell pellet was resuspended in 10 ml ice-cold cell lysis buffer (5 mM pH8.0 PIPES buffer, 85 mM potassium chloride, 0.5% Nonidet P40) containing protease inhibitors (cOmplete, EDTA-free, Roche, Indianapolis, IN, USA) for 10 minutes at 4 °C. Nuclei were released by douncing with a Wheaton homogenizer pestle B. The crude nuclear extract was collected by centrifugation at 500 x g (1680 rpm) for 5 min at 4 °C, resuspended in 2 ml ice–cold nuclear lysis buffer (50 mM pH8.1 Tris.HCl, 10 mM EDTA, 1 % SDS with protease inhibitors) and incubated for 20 minutes at 4 °C. After adding 1 ml ice-cold IP dilution buffer (0.01 % SDS, 1.1 % TritonX-100, 1.2 mM pH 8 EDTA.Na2, 16.7 mM pH8 Tris.HCl, 167 mM NaCl and protease inhibitors) and 0.3 g acid-washed glass beads (Sigma-Aldrich, St. Louis, MO, USA) to the nuclear extract, the chromatin was sheared to 200-1000 bp using a Misonix Sonicator 3000 (Misonix, Inc. Farmingdale, NY, USA). Sonication was performed on ice water with 8 pulses of 30 seconds at 30 second intervals. Thereafter, the cell debris were removed by centrifugation at 16000 x g (13000 rpm) for 10 min at 4 °C. Input DNA was prepared in an identical manner from non-transfected cells.

The sonicated, fixed chromatin was precleared by incubation with preblocked magnetic Protein G coupled Dynabeads (Invitrogen, Carlsbad, CA, USA) overnight at 4 °C on a rotating wheel. Subsequently, the chromatin was divided into three aliquots of 850 μg. IP samples were incubated with 8.5 μg anti-V5 tag monoclonal antibody (Inivitrogen, Carlsbad, CA, USA)

prebound to Dynabeads overnight at 4 °C on a rotating wheel. The beads were washed three times with low salt buffer (0.1 % SDS, 1 % Trition, 2 mM EDTA, 20 mM pH 8 Tris, 150 mM NaCl) three times with high salt buffer (0.1 % SDS, 1 % Trition, 2 mM EDTA, 20 mM pH 8 Tris, 500 mM NaCl) and finally twice with LiCl buffer (10 mM pH 8.1 Tris, 1mM EDTA, 0.25M LiCl, 1 % NP40, 1 % sodium deoxycholate) with incubation at room temperature on a rotating wheel for 5 min respectively. The beads were incubated twice on a rotating wheel at room temperature for 20 min in 200 μl elution buffer (0.1 M $NaHCO_3$ and 1% SDS) to recover the immunoprecipitated DNA. Cross-links were dissociated by incubation at 65 °C overnight. DNA was purified by phenol-chloroform extraction and ethanol precipitation.

100 ng immunoprecipitated DNA and 300 ng of input DNA were used to make libraries with the Genomic DNA sample preparation kit (Illumina, San Diego, CA, USA) according to the manufacturer's protocol. Adapter-ligated DNA of 200 ± 25 bp range was excised from the gel before PCR amplification. Input chromatin was prepared from two biological replicates and IP samples were prepared from three biological replicates. ChIP-seq libraries were sequenced on an Illumina GA1 instrument with either 25 or 36 bp read lengths. Reads were generated using the Illumina pipeline software 0.3.0. All ChIP-seq reads were trimmed to 25 bp prior to mapping to the *D. melanogaster* genome (FlyBase release 5 with no Uextra). The sequence reads from all biological replicates were pooled prior to mapping using Bowtie 0.12.7 (Langmead et al., 2009) accepting only uniquely mapped reads with no more than 2 mismatches (-m1 –v2). Duplicate reads were removed from the libraries before peak calling with the Picard tool MarkDuplicates v1.95 (http://picard.sourceforge.net). The WTD method of peak calling from the ChIP-seq analysis program SPP v1.11 was used to call peaks with an FDR of 0.01 (Kharchenko et al., 2008). IP and input reads were loaded into SPP and anomalous reads due to localized regions of

extremely high read count were removed with the command *remove.local.tag.anomalies*. Broader peak regions of enrichment surrounding the predicted binding site were added to create the final peak coordinates using the command *add.broad.peak.regions*. In order to calculate ChIP-seq signal across the genome for use in producing a gene-level occupancy score (see below), IP or input reads were counted in non-overlapping consecutive 500 bp intervals across the genome with HTSeq v0.5.1p2 (Anders et al., 2014). Read counts were depth-normalized and the log2 Fold Change [(IP + 1)/(input +1)] was calculated for all bins.

de novo motif analysis

The binding positions reported for the DSX^M and DSX^F proteins from the SPP ChIP-seq analysis (above) were sorted by descending binding scores and the top 1000 scoring sites were selected for further analysis. 200 bp of DNA on either side of the identified binding position were used to search for enriched DNA sites using MEME-ChIP (Machanick and Bailey, 2011). Comparison of the position weight matrix for the biochemically determined DSX binding sequence (Yi and Zarkower, 1999) to the *de novo* site analysis from MEME for either DSX isoform was performed with TOMTOM (Gupta et al., 2007).

Gene-level occupancy scores

Gene level occupancy scores were calculated by summing the log2 fold change of (DSX occupancy signal)/(Control signal) in 500 bp bins under all called peak regions within the gene body plus 1 kb upstream of the TSS (adapted from (Ouyang et al., 2009)). We picked the window by analyzing the relationship between occupied regions and gene features (Figure S2C) and the analysis showed that the DSX-bound sites were primarily observed within 1 kb

preceding the TSS (with maximum at 0.6 kb before TSS) and decreased in density throughout the gene body. For genes with no called peaks, the gene level occupancy score was computed as the average log2 Fold Change of (DSX occupancy signal)/(Control signal) over 500 bp bins in the gene body plus 1 kb upstream. Genes were sorted in non-increasing order of gene level occupancy scores (genes with peaks ranked first and then genes without peaks followed). The ranks were normalized by dividing the gene ranks by the total number of genes.

Gene-Level DSX PWM Score

The position weight matrix (PWM) for DSX sequence binding was composed of the position nucleotide percentages reported for DSX protein (Yi and Zarkower, 1999). The PWM was converted into the JASPAR format and was used to search the *D. melanogaster* genome (FlyBase release 5) for sequences matching the PWM with the Bio.Site.search_pwm method module in BioPython (Cock et al., 2009). 556,628 sequences with any relationship to the DSX PWM were identified using this method. Each sequence was assigned a score calculated by summing the log odds for each position. Scores ranged from 0.000103 to 18.84745 (Table S2).

The gene-level DSX PWM score (Table S1) was based on the number of DSX binding sequences at a gene as well as the PWM score of each site. For a gene g, let S(g) be the set of DSX binding sequences within gene body plus 1kb upstream. We computed the probability that at least one binding event occurs in S(g) and used this as the gene level DSX PWM score, assuming that binding events are independent and that the probability of binding to a sequence is $(1+\varepsilon)^{w_i-W}$ where w_i is the PWM score of i-th sequence in S(g) multiplied by 10, W = 189 (rounded up from the maximum PWM score of a sequence multiplied by 10) and an adjusting parameter $\varepsilon = 0.03$. Gene level DSX PWM score is then defined as follows.

$$PWMscore(g) = 1 - \prod_{m_i \in S(g)} (1 - (1+\varepsilon)^{w_i - W})$$

where m_i is i-th sequence in S(g).

Conservation of DSX binding sequences and gene-level DSX conservation index (CI) score

The genomes of *D. simulans, D. sechellia, D. yakuba, D. erecta, D. ficusphila, D. eugracilis, D. biarmipes, D. takahashii, D.elegans, D. rhopaloa, D. kikkawai, D. ananassae, D. bipectinata, D. pseudoobscura, D. persimilis, D. willistoni, D. mojavensis, D. virilis* and *D. grimshawi* (Adams et al., 2000; Chen et al., 2014; Drosophila 12 Genomes et al., 2007; Richards et al., 2005) were searched for sequences that relate to DSX's binding sequence position weight matrix as described above for *D. melanogaster*. All identified sequences were associated with genes according to the identity of the nearest first coding exon using BEDTools v. 2.16.2 (Quinlan and Hall, 2010). For genes with multiple transcripts with different first protein-coding exons, only the most proximal first-coding exon was used. The positions of all first coding exons in each species were identified by aligning first coding exons from *D. melanogaster* (FlyBase annotation version 5.46) using liftover chain files. To create liftover chain files, whole-genome alignments between *D. melanogaster* and each other Drosophila species were performed using lastz (Harris, 2007) and executables from the UCSC Genome Browser (Meyer et al., 2013) according to a protocol on the UCSC user guide. Briefly, genomic sequences from each non-melanogaster species were split into 5 MB segments with the faSplit executable (parameters: size -oneFile 5000000 -extra=10000), and pairwise alignment was performed against *D. melanogaster* with lastz (parameters: --masking=50 --hspthresh=2200 --ydrop=3400 --gappedthresh=4000 --inner=2000). These alignments were converted to Pattern Space Layout (PSL) format and lifted to chromosomes with the lavToPsl and liftUp executables. Then, these

PSL alignments were chained with the axtChain executable (parameters: -linearGap=medium -psl), combined with the chainMergeSort and chainSplit executables, and converted to alignment nets with the chainNet executable. Based on alignment nets, liftOver chain files that convert annotations from *D. melanogaster* to other species were created with the netChainSubset executable.

D. melanogaster sequences with positive PWM scores located within a protein-coding *D. melanogaster* gene body plus 1 kb upstream excepting those in coding sequence or located on chrU, chrUextra or chrM (173,775 in total, Table S2) were used to search for orthologous sequences among those sequences that were associated with the same gene in each of the remaining 19 genomes. A sequence was considered orthologous if the edit distance of the largely invariant nucleotides at position 4-10 was ≤ 1 and the position difference relative to the first coding exon was less than 2 kb. A conservation index (CI) score for each sequence (Table S2) was computed by summing the substitution/site distance associated with each species in which the sequence was identified. The evolutionary distances between *D. melanogaster* and 19 other Drosophila species (Figure S1) are expressed in units of substitutions per synonymous site (ss) as defined in (Chen et al., 2014).

A CI score for each gene (Table S1) was computed either by summation of CI scores >90th percentile for all sequences associated to the gene (gene body plus 1 kb upstream excluding coding sequence) or by taking the maximum CI score associated with a gene. The two methods yielded similar results (Spearman's rho 0.7861143). The summation method was chosen as it provided higher resolution gene-level CI scores since many genes could have identical maximum CI values. The 90 percentile threshold chosen for the summation method was chosen based on

the observation that the break point in normalized CI score is present in the range of 80-90 percentile.

As a null model of DSX conservation, 100 random motifs were generated by randomly shuffling the 13 positions of the DSX PWM. For each of 100 shuffled PWMs, we identified sites in all species with positive PWM scores and calculated the site-level and gene-level CI scores using the same method by restricting the edit distance of the corresponding invariant positions in shuffled motifs ≤ 1.

The normalized site level CI scores (Figure 2A) were calculated by subtracting median CI score of shuffled motifs from DSX CI score as follows: DSX sites are sorted and divided into 1000 bins of equal number of sites and moving median of CI scores is calculated in each window of all 10 consecutive bins. The moving median of CI scores of all 100 shuffled motifs is also obtained by considering the same number of sites as DSX motifs in each window. The median of CI scores of the 100 shuffled motifs is then computed in each window and used to subtract from DSX CI scores.

Conservation Analysis using PhastCons

PhastCons first performs multiple alignments over 15 species and uses two-state phylogenetic hidden Markov model (phylo-HMM) to predict conserved elements. PhastCons scores (Felsenstein and Churchill, 1996) were downloaded from UCSC (WIB files from http://hgdownload.cse.ucsc.edu/gbdb/dm3/multiz15way/wib; SQL table dump from http://hgdownload.cse.ucsc.edu/goldenPath/dm3/database/phastCons15way.txt.gz). The UCSC program hgWiggle (from http://hgdownload.soe.ucsc.edu/admin/exe/linux.x86_64) was used to convert each WIB file into a WIG file for each chromosome containing phastCons scores for the

15-way multiple alignment performed by UCSC (see http://genome.ucsc.edu/cgi-bin/hgTrackUi?g=multiz15way for processing details). WIG files were manually edited to remove duplicate lines that prevented conversion to bigWig format, and the edited files were converted to bigWig using the UCSC program wigToBigWig. Data from all chromosomes were then concatenated into a single file with the UCSC program bigWigCat. A site Level PhastCons score is obtained by calculating the average of PhastCons scores of all positions with the UCSC program bigWigAverageOverBed.

Conservation of DSX DNA binding domain and splicing patterns

Exons coding for DSX were identified in 19 non-melanogaster species by liftover of the *Drosophila melanogaster* DSX-encoding exons as described above. Exons were translated *in silico* using ExPASy (Artimo et al., 2012). The DM DNA binding domain was defined as *D. melanogaster* DSX amino acids 39-105 (Zhang et al., 2006). Multiple species alignment of DSX DM domain protein sequence was performed with Clustal Omega 1.1.0 (Sievers et al., 2011). A single nucleotide (G) deletion was identified in the DM domain of the DSX coding sequence in the genome of *D. mojavensis* between nucleotide positions 23339379 - 2333980 of scaffold_6540 (Genbank accession NW_001979112.1). This deletion may be due to an assembly artifact as the deleted G was present in RNA-seq data from *D. mojavensis* (Chen et al., 2014).

DMRT1 Orthologs

952 *D. melanogaster* orthologs of mouse DMRT1 targets (Murphy et al., 2010) were identified by converting the 1,439 gene names to Ensembl IDs using the Jackson Labs conversion tool (Blake et al., 2014). These Ensembl IDs were uploaded into Ensembl Biomart

(Flicek et al., 2013), and orthologous *D. melanogaster* genes were obtained using the multispecies comparison tool with Ensembl 73 Genes.

Occupancy score clustering and statistical analysis

The ranked occupancy scores for all genes annotated in FlyBase 5.46 in all six occupancy experiments were clustered using the kmeans package in R (R Core Team, 2013). The Kruskal-Wallis test was performed in R (kruskal.test function) and used to test for significant differences in the distributions of gene-level DSX PWM scores and gene-level conservation for genes in each cluster. The hypergeometric test was used to test for significant enrichment of *D. melanogaster* orthologs of mouse DMRT1 targets among the genes in each occupancy cluster.

ChIP-seq correlations with multiple independent occupancy data sets

Called peaks for 255 available ChIP-chip and ChIP-seq experiments annotated in Slattery et al's supplemental table S1 (Slattery et al., 2014) as well as those available in other modENCODE accessions were downloaded in GFF, BED, plain text tables, or tarballs from either individual GEO entries, authors' websites, or the modENCODE FTP site (Table S6). Files for called peaks were converted to a uniform BED3 format, and peaks from the Furlong group and BDTNP project were lifted over to the dm3 assembly. Peaks from the occupancy experiments from this study were also included for a total of 261 sets of called peaks. In each file, features that overlapped by at least 1 bp were merged using BEDTools v2.19.0. Since some experiments only included euchromatic chromosomes, for consistency all files were filtered to only retain peaks from chromosomes 2L, 2R, 3L, 3R, 4, and X. Details on data acquisition and processing can be found in Table S6.

For each pairwise comparison, one file was arbitrarily set as the query and one file as the reference. P-values for each peak in the query were calculated following the IntervalStats method of (Chikina and Troyanskaya, 2012), representing that peak's overall proximity to the reference. The similarity of the query to the reference was then summarized in a single number by taking the fraction of all features in the query with p-values < 0.05. Since this metric is not symmetric, the fraction of features with p-values < 0.05 was also calculated after swapping the query and reference. The final result is a 261 x 261 similarity matrix of pairwise comparisons with each value representing the fraction of all peaks with P < 0.05 in the query (Figure S3).

Defining DSX-occupied genes

3,717 genes were defined as being occupied by DSX in all experimental data sets by taking the union of all genes in the 90th percentile of gene-level occupancy scores from each individual occupancy data set (Table S1). The 90th percentile cutoff was selected for use in the analysis following examining the relationship between gene-level occupancy score and gene-level PWM score. The best break point in these plots were where CI score was 80-90 percentile. 2,668 genes were defined as being occupied by DSX in fat body samples by taking the union of the genes in the 90th percentile of the three fat body occupancy data sets.

Analysis of occupancy and conserved motifs relative to gene features

Using all genes > 1 kb annotated in FlyBase 5.46, we binned loci into five regions: upstream (1.5 kb upstream of 5'-most promoter, 1-bp bins); 5' (500 bp downstream of promoters, 1-bp bins); gene body (1000 bins; bin size varies); 3' (500 bp, 1-bp bins); and downstream

(1.5kb, 1-bp bins). The number of DSX peaks (union of peaks from all occupancy experiments; median size 1kb) were enumerated in each bin using metaseq v0.5 (Dale et al., 2014) and averaged across all genes. Values were then normalized by subtracting the minimum and dividing by the maximum.

GOTerm Analyses

Enrichment of gene ontology terms (Table S5) was identified using the Cytoscape app BINGO 3.0.2 (Maere et al., 2005). The genes from each individual occupancy cluster were used as the input dataset, and the total *D. melanogaster* gene set was used as the background file. p-values returned by BINGO are corrected for multiple testing using the Bonferroni method. We considered adjusted p-values < 0.001 as a significant enrichment.

RNA-seq

Fat body tissue was dissected from age-matched adult flies of the genotypes w^{1118}; $tra2^{ts2}$/ $tra2^{ts1}$ (experimental) or w^{1118} (control for dsx^{F}->dsx^{M} experiments); for dsx^{M}->dsx^{F} experiments the genotypes were: *y^{1} w^{*}; P{w^{+mc}=UAS-Tra.F}20J7; P{w^{+mc}=tubP-GAL80^{ts}}7/P{w^{+mc}=tubP-*GAL4*}LL7* (experimental) or *P{w^{+mc}=tubP-GAL80^{ts}}7/P{w^{+mc}=tubP-*GAL4*}LL7* (control). All samples were raised at 18°C until 5 days after eclosion when adults were shifted to either 29°C (for *$tra2^{ts}$*) or 30°C (for *UAS-TraF*) for 0, 12, or 24 hours. Total RNA was extracted from fat body dissected at room temperature (placed on ice after 30 minutes) using TRIzol Reagent following manufacturer's protocol (Ambion Life Technologies, Carlsbad, CA, USA). Purified RNA was treated with DNAse I following manufacturer's protocol (New England Biolabs, Ipswich, MA, USA) and purified again using phenol:chloroform extraction followed by ethanol

precipitation. Duplicate RNA-seq libraries were constructed from 200ng total RNA from independent dissection of each sample using the TruSeq RNA Sample Preparation v2 high-throughput (HT) protocol (Illumina, San Diego, CA, USA, 2011). Libraries were sequenced on the HiSeq 2000 machine following a 76 bp single-end protocol (Illumina, San Diego, CA, USA).

Reads were generated using the Illumina pipeline software 12.4.2 for all samples excluding control male t=24hr replicate 1 which used pipeline 1.13.48 (re-sequenced due to poor initial sequence quality). Reads passing the Illumina chastity filter were mapped to the *D. melanogaster* genome and assigned to gene models using Tophat 1.4.1 (Trapnell et al., 2009) with a gtf file provided (-G, FlyBase r5.46, see below) and default settings except for the following; minimum intron length was set to 42bp (-i 42) and the maximum multihits was set to 1 (-g 1). Transcript abundance was determined using Cufflinks 2.1.1 (Trapnell et al., 2013) with maximum bundle fragments set to 10,000,000 (--max-bundle-frags 10000000) due to high read density at the *Yp* loci, and upper quartile normalization was used (-N).

To generate a gtf file for Tophat and Cufflinks analyses, the FlyBase GFF annotations (release 5.46) were downloaded from FlyBase as a GFF3 format file. This file was filtered to remove any features on chromosomes Uextra or dmel_mitochondrion_genome as well as the following feature types: enhancer, regulatory_region, exon_junction, rescue_fragment, sequence_variant, pcr_product, point_mutation, orthologous_region, TF_binding_site, protein, chromosome, uncharacterized_change_in_nucleotide_sequence, origin_of_replication, chromosome_band, tandem_repeat, insulator, polyA_site, deletion, BAC_cloned_genomic_insert, complex_substitution, RNAi_reagent, transposable_element_insertion_site, repeat_region, oligonucleotide, breakpoint, transposable_element, chromosome_arm, protein_binding_site, orthologous_to, silencer, region,

insertion_site, mature_peptide, DNA_motif, syntenic_region. A leading "chr" was prepended to each chromosome name for consistency with the genomic assembly sequence files used. The filtered GFF file was imported into a sqlite3 database using gffutils (https://github.com/daler/gffutils), which represents the hierarchical relationships between features as defined in GFF files. For each gene, the "child" transcripts were retrieved from the database, and for each transcript, each child that was either an exon or CDS was retrieved. For each of these exon and CDS features, the gene ID, gene name, transcript ID, and transcript type information were attached to the feature, and it was exported as a GTF format line. The resulting GTF file of exon and CDS features was then run through the gffread program (part of the cufflinks suite) as the command "gffread -E $infile -T -F -o- > $outfile" in order to confirm that the file contained no errors that would prevent downstream use by Cufflinks (Trapnell et al., 2012).

Background expression levels were estimated based on reads in intergenic space (Zhang et al., 2010). Genomic regions that are not located within an annotated gene (FlyBase 5.46), nor within +/- 500 bp flanking an annotated gene, were binned into 199 bp windows (= median of all *D. melanogaster* exons), and FPKM values for these intergenic bins were calculated using the Tophat/Cufflinks parameters used for genes. To prevent loss of mapping between bins, the original intergenic bins were shifted by 100bp and any bin entering a non-intergenic space was removed. The median expression value for all intergenic bins was 1.84839375, and all experimental FPKM values at or below this cutoff were converted to zero. Further, all genes with FPKM=0 in all experimental and control conditions were removed from further analyses. After background correction, k-means clustering of FPKM values was performed using the kmeans package in R. The optimal k value (k=9) was determined by maximizing both k and

average silhouette width. Counts of dsx^M and dsx^F splice junctions were obtained using Spanki 0.4.2 (Sturgill et al., 2013).

GEO Accession Numbers

DSX occupancy data (ChIP-seq, DamID-seq, DamID-array) and RNA-seq data are available under GEO series accession GSE49480. Probe sequence, probe position information, and array details are available under GEO accession GPL10639.

Results

DSX occupancy

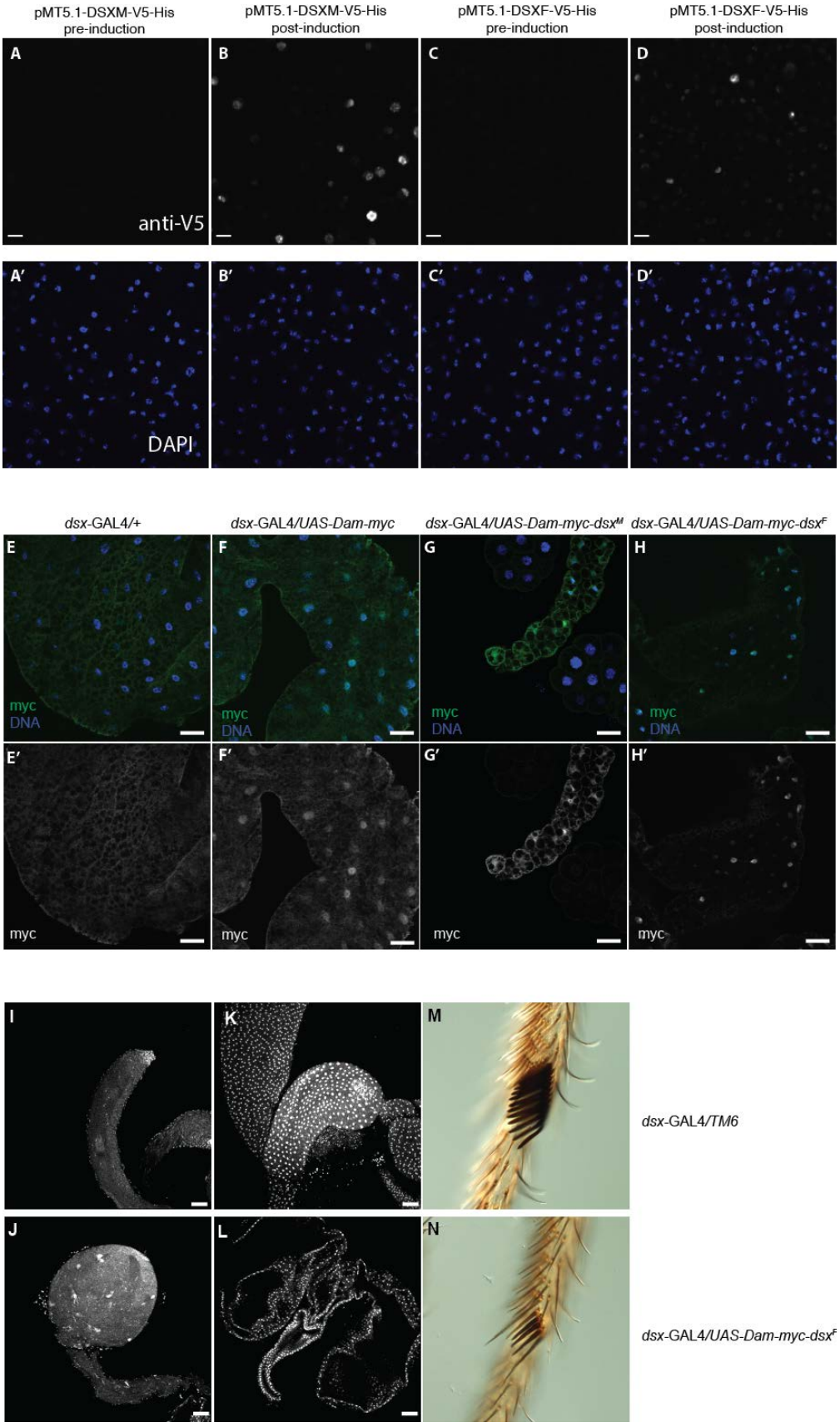
To empirically determine where DSX binds in the *D. melanogaster* genome, we performed chromatin immunoprecipitation followed by sequencing (ChIP-seq) on S2 cells expressing tagged DSX^M or DSX^F. We also performed DSX^M or DSX^F DNA adenine methyltransferase identification (DamID) on adult ovary and adult female and male fat body in transgenic flies followed by either sequencing (DamID-seq) or hybridization to microarrays (DamID-chip). In particular, we chose the adult fat body and ovary since *dsx* is known to play a role in maintaining sexual dimorphic gene expression in both organs (Burtis et al., 1991; Coschigano and Wensink, 1993; and MVD, unpublished). We confirmed nuclear expression of tagged forms of DSX by immunohistochemistry of S2 cells and Drosophila tissues. When we drove expression under the control of the *dsx* promoter (*dsx*-GAL4), we detected nuclear localization of the tagged DSX proteins. We also observed nuclear localization of the unfused

DAM protein revealing that it too localizes to the nucleus where it can freely methylate the genome in regions of open chromatin as a background control (Figure 2.1). Further, expression of *Dam-dsx*^{F} in males using *dsx*-GAL4 induced feminization of appropriate tissues such as the sex combs, reproductive tract, and gonads, indicating that these constructs are functional (Figure 2.1, over expression of *Dam-dsx*^{M} was lethal). DamID experiments were conducted using the low basal expression of Dam-*dsx* in the absence of a GAL4 driver to avoid known toxicity associated with DAM methylase expression ensuring that our DSX occupancy profiles were not subject to artifacts due to DSX over expression. For all samples, we explored the continuous distribution of DSX occupancy using background subtracted values to control for general chromatin accessibility. We identified "peaks" of occupancy using a stringent 1% FDR cutoff (Greil et al., 2006; Southall and Brand, 2007). Data tracks and called peaks are available at the Gene Expression Omnibus (GEO) under accession GSE49480.

The first step in the occupancy analysis was at the level of contiguously occupied regions, or peaks. We expected DSX occupancy in regions near known DSX target genes (Figure 2.2AB). Indeed, the divergently transcribed *Yp1* and *Yp2* loci showed strong DSX occupancy in the fat body and ovary where these genes are expressed at high levels along with weak occupancy in S2 cells. In contrast, the *bab1* locus showed strong DSX occupancy in all samples. While we observed occupancy at the previously identified *Yp1/2* intergenic and *bab1* intronic DSX response elements, we also found a strongly occupied region upstream of *bab1* that may represent an additional DSX-dependent enhancer.

We next wanted to associate the DSX binding sites in the genome with the closest relevant genes and generate a score for DSX occupancy for all genes in the genome. *Yp1/2* and *bab1* appear to be typical examples of DSX occupancy patterns as we found a strong preference

Figure 2.1. Expression of Dam-DSX fusion protein and resultant phenotypes. S2 cells carrying the pMT5.1-DSXM-V5-His construct before (A, A') and after (B, B') 60 hour induction. S2 cells carrying the pMT5.1-DSXF-V5-His construct before (C, C') and after (D, D') 60 hour induction. Scale bar = 10μm. A, B, C, and D are the V5 channel (white), and A', B', C', and D' are the DAPI channel (blue). Third instar larval fatbody from *dsx*-GAL4/+ (E, E') *dsx*-GAL4/ *UAS-Dam-myc* (F, F') *dsx*-GAL4/*UAS-Dam-myc-dsx*^{M} (G, G') and *dsx*-GAL4/*UAS-Dam-myc-dsx*^{F} (H, H'). E,F,G, and H are merged images of anti-myc (green) and DAPI (blue). E', F', G' and H' is a split of only the anti-myc signal (white). Testes (I,J) and ejaculatory ducts (K,L) were dissected and stained with DAPI. Light microscopy images of sex combs from control *dsx*-GAL4/*TM6* (M) and *dsx*-GAL4/*UAS-Dam-myc-dsx*^{F} (N). Scale bar = 50μm. The Dam fusion proteins include the myc epitope incorporated at the C-terminus of the Dam coding sequence such that Dam fusion proteins can be detected with anti-myc antibodies.


pMT5.1-DSXM-V5-His
pre-induction
pMT5.1-DSXM-V5-His
post-induction
pMT5.1-DSXF-V5-His
pre-induction
pMT5.1-DSXF-V5-His
post-induction
A
B
C
D
anti-V5
A′
B′
C′
D′
DAPI
dsx-GAL4/+
dsx-GAL4/UAS-Dam-myc
dsx-GAL4/UAS-Dam-myc-dsxM
dsx-GAL4/UAS-Dam-myc-dsxF
E
F
G
H
myc
DNA
E′
F′
G′
H′
myc
I
K
M
J
L
N
dsx-GAL4/TM6
dsx-GAL4/UAS-Dam-myc-dsxF

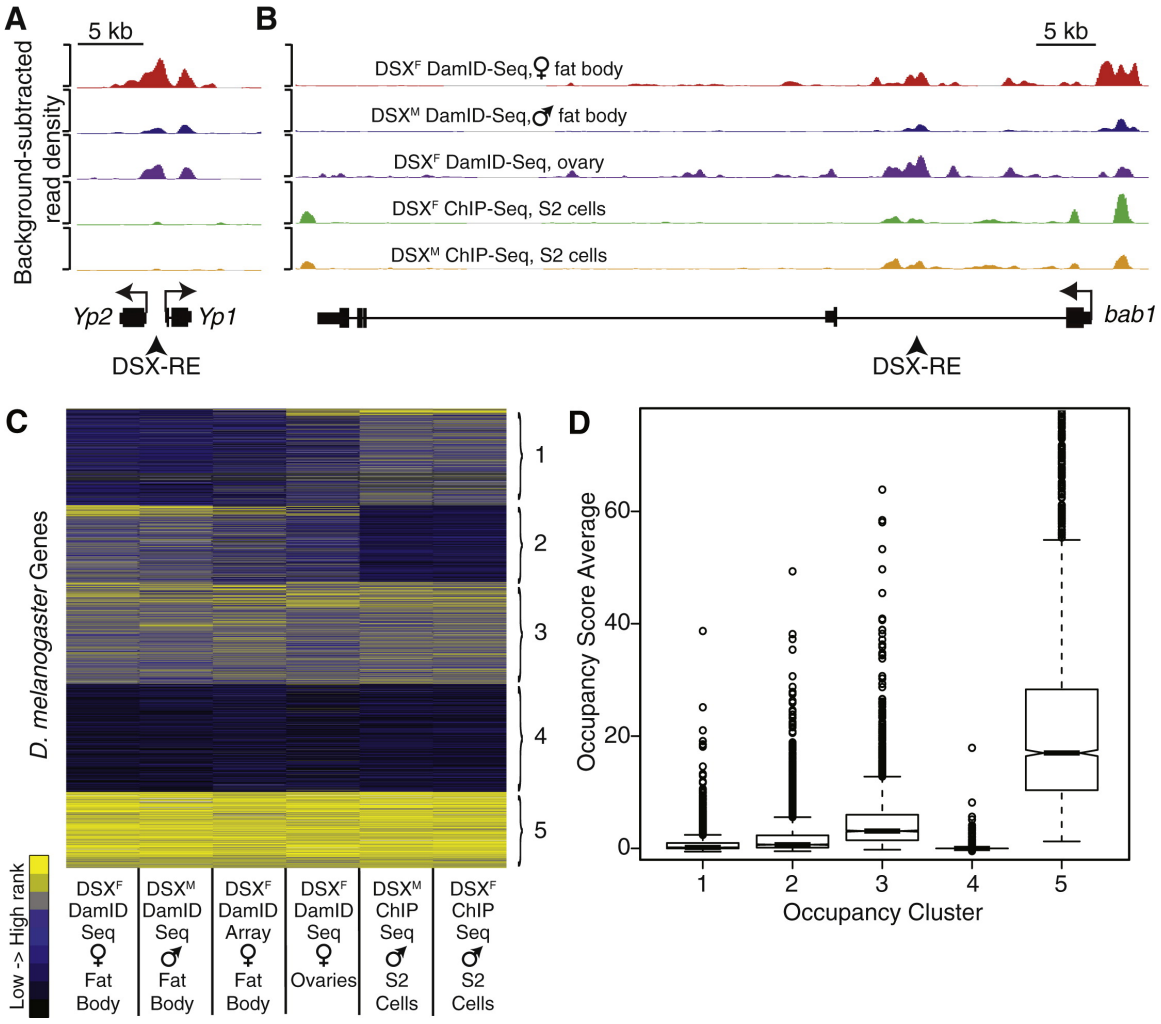
for occupancy within the first 1 kb upstream of the start of transcription and near the 5' end of gene bodies (Figure 2.2). Because of this pattern, we assigned DSX peaks to genes either by using peaks occurring within the gene body plus 1 kb upstream or by using a fixed 2 kb window centered on the annotated transcription start site (fixed-range method). In both cases this peak-to-gene definition limits artificial contributions of nearby upstream genes; however, the fixed-range method uncouples gene length from occupancy analysis yet misses binding at sites such as downstream intronic enhancers. The gene body plus 1kb definition captures these enhancers, but biases occupancy toward longer genes. We elected to use "gene plus 1 kb" definition as this captured genes with intronic enhancers, such as *bab1,* however both methods produced a largely overlapping list of occupied genes.

We determined occupancy strength at each gene by the strongest peak at a gene (peakmax method) or the sum of all peaks at a gene (peaksum method) (Table S1 doi: 10.1016/j.devcel.2014.11.021). Here, we elected not to normalize for gene length, since it introduced a strong bias against long genes, such as *bab1,* with relatively few discrete but strong DSX binding regions. The two occupancy strength methods produced similar occupancy ranks (Spearman's rank correlation coefficient > 0.9) and we chose to use the peaksum method as it appropriately yielded higher occupancy strength values for genes with multiple, strong binding events.

There are many ways to examine the relationships between DSX occupancy patterns at different loci. Supervised (k-means, where k = 5) clustering of the ranked occupancy of all genes allowed us to identify groups that exhibited a similar pattern of occupancy across our different occupancy data sets (Figure 2.2C, Table S1 doi: 10.1016/j.devcel.2014.11.021). This analysis yielded clusters of DSX occupancy patterns among genes that exhibit no to very low occupancy (cluster 4), tissue non-specific occupancy (clusters 3 and 5), and tissue- and/or

Figure 2.2. DSX occupancy and binding sites.

(A-B) Scaled read density plots (background subtracted) from five replicated occupancy experiments (as labeled) along the genome for (A) the *Yolk protein 1 & 2* (*Yp1*, *Yp2*), and (B) the *bric-a-brac 1* (*bab1*) loci (arbitrary scale). FlyBase gene models showing transcription start sites (bent arrows), coding exons (thick rectangles), non-coding regions (thin rectangles), introns (lines), and positions of known DSX response elements. (C) Heatmap of k-means clustering of background-subtracted, ranked occupancy score for *D. melanogaster* genes (optimal k value (k = 5, see Supplemental Experimental Procedures). (D) Box plots of gene-level occupancy scores averaged from 6 occupancy data sets in each occupancy cluster from (C). In all occupancy experiments, replicate preparations of a given sample type showed excellent reproducibility (spearman rho > 0.9, data not shown). Color scale for ranked occupancy scores is given on the left.


A
5 kb
Background-subtracted read density
Yp2
Yp1
DSX-RE
B
5 kb
DSXF DamID-Seq,♀ fat body
DSXM DamID-Seq,♂ fat body
DSXF DamID-Seq, ovary
DSXF ChIP-Seq, S2 cells
DSXM ChIP-Seq, S2 cells
bab1
DSX-RE
C
D. melanogaster Genes
1
2
3
4
5
Low -> High rank
DSXF DamID Seq ♀ Fat Body
DSXM DamID Seq ♂ Fat Body
DSXF DamID Array ♀ Fat Body
DSXF DamID Seq ♀ Ovaries
DSXM ChIP Seq ♂ S2 Cells
DSXF ChIP Seq ♂ S2 Cells
D
Occupancy Score Average
0
20
40
60
1
2
3
4
5
Occupancy Cluster

technique-specific occupancy (clusters 1 and 2). In this particular analysis, the *bona fide* DSX target *bab1* was in cluster 5 while the *Yp1/2* loci were in cluster 3 due to modest occupancy in S2 cells. Genes ranking in the top 10% of occupancy were almost exclusively in cluster 5. Genes outside of clusters 3 and 5 contained genes with low overall absolute occupancy values although there were a few genes with strong occupancy in each cluster (Figure 2.2D).

Interestingly, the DSX^F and DSX^M proteins showed similar occupancy patterns (Figure 2.2C), suggesting that the sex-specific effector domains and sex-biased chromatin environments had a negligible impact on where DSX binds. However, there are transcriptional "hotspots" that are known to bind a host of different factors (Negre et al., 2011). To determine whether the tissue non-specific occupancy and common DSX^F and DSX^M patterns we observed were due to non-specific DSX binding at accessible chromatin or at hotspots, we correlated our 6 occupancy data sets with 255 occupancy experiments for different DNA-associated proteins (see Materials and Methods, doi: 10.1016/j.devcel.2014.11.021). DSX^F and DSX^M occupancy patterns were more similar to one another than they were to any other assayed factor indicating that our DSX occupancy patterns are not simply due to chromatin accessibility or binding at hotspots (Figure 2.4). Additionally, deliberately removing genes associated with hotspots prior to analysis did not influence the overall structure of the occupancy patterns (not shown).

We conclude that the strongest *in vivo* binding of DSX in the genome occurs in a largely tissue non-specific manner. This observation focused our attention on genes with strong occupancy in each sample, but there were certainly genes with tissue-specific or isoform-specific occupancy patterns that may prove to be extremely interesting for future work. Since DSX has dramatically different genetic roles in different tissues and between the sexes, focusing on this

set of genes allowed us to address a previously unexplored question of how DSX integrates with other tissue-specific factors rather than regulation simply by where DSX binds in the genome.

Sequence analysis of DSX binding sites

We hypothesized that the observed occupancy pattern would be primarily due to direct binding by DSX while other contacts might be indirect due to 3D structures such as looping. One simple prediction of this hypothesis is that there should be a correlation between regions in the genome occupied by DSX and those that contain a DSX binding site. DSX DNA binding-specificity has been biochemically-defined in *D. melanogaster* (Erdman et al., 1996; Murphy et al., 2007; Yi and Zarkower, 1999). We found significant enrichment for sequences matching the DSX motif position weight matrix (PWM) in regions of occupancy both by *de novo* motif finding under occupied regions from ChIP-seq ($p < 0.01$), and by scanning the entire genome with the PWM for DSX ($p < 0.01$; Fisher's Exact Test).

A major problem with transcription factor binding site studies is that these short sequences are quite common in the genome and factors bind these sites in both functional and non-functional contexts (Fisher et al., 2012). Thus, the presence of an occupied DSX site in the genome is not *de facto* evidence of function. To enhance our ability to predict functional DSX binding sites, we used comparative genomics to analyze the conservation of DSX binding sites among 20 species of Drosophila with sequenced genomes (Adams et al., 2000; Chen et al., 2014; Drosophila 12 Genomes et al., 2007; Richards et al., 2005). While conservation is not always predictive of function (Villar et al., 2014), and some non-conserved sites may be interesting as species-specific targets, sites conserved over evolutionary timescales are likely to function in the regulation of the vast array of genes showing sex-biased expression in the genus (Chen et al.,

Figure 2.3. Conservation of the DSX DNA binding domain and sex-specific splicing. (A) Diagram of the DSX^M and DSX^F proteins (above) and amino acid sequence alignment of the DSX DNA binding domain from 20 Drosophila species (below). Cysteine and histidine residues in the Zn-binding site are highlighted in tan. The evolutionary distance from *D. melanogaster* is indicated in substitutions/site (ss) (Chen et al., 2014). Color-coding of *D. melanogaster* amino acids represent mutations that do not affect DSX activity (green), partially affect activity (orange), or impair activity (red) (adapted from (Zhang et al., 2006)). (B) Bar graphs representing the percentage of *dsx* splicing events resulting in production of female (red) or male (blue) isoform from RNA-seq data obtained from adult females (F) or males (M) from 7 Drosophila species (Chen et al., 2014). (C) The normalized (% of maximum average occupancy) distribution of DSX occupancy values along a generic gene model using +1.5Kb upstream of transcription start (bent arrow), the gene body (rectangle), where the first 0.5Kb and last 0.5Kb are at base level, and the middle 0.5Kb is scaled from all gene models, and the -1.5Kb downstream region are shown.

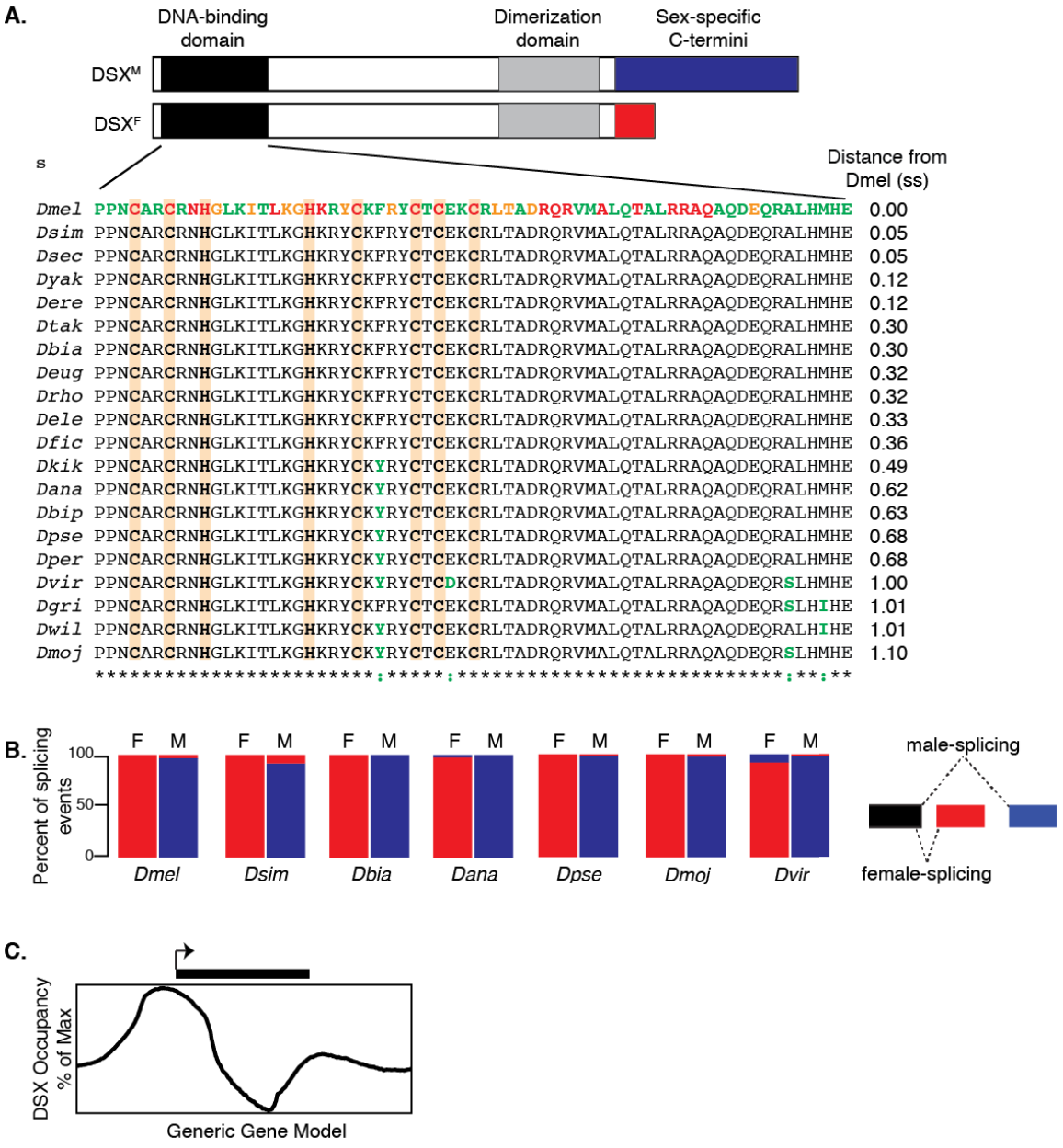
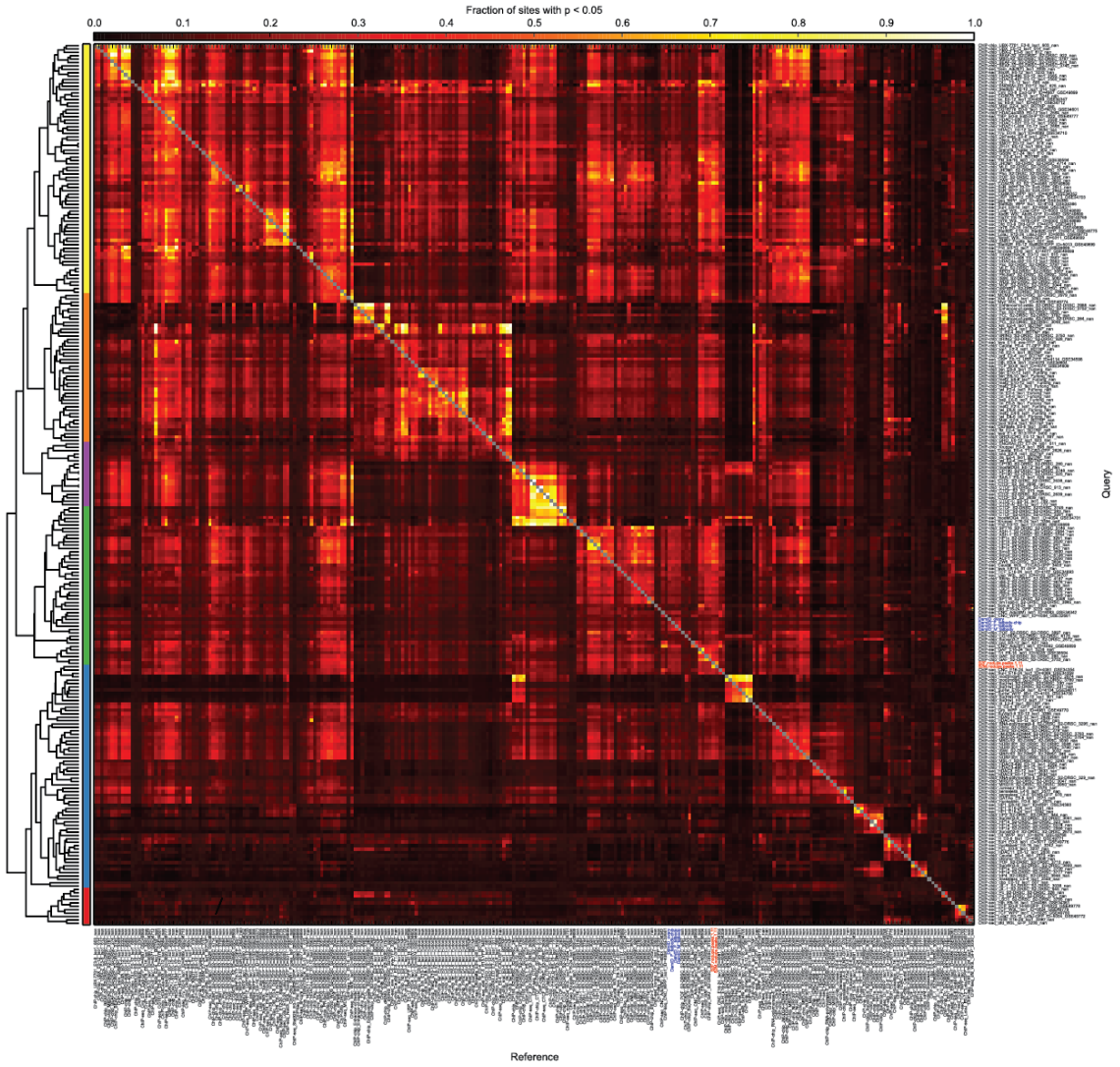

A.
DNA-binding domain
Dimerization domain
Sex-specific C-termini
DSXM
DSXF
s
Distance from Dmel (ss)
Dmel PPNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.00
Dsim PPNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.05
Dsec PPNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.05
Dyak PPNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.12
Dere PPNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.12
Dtak PPNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.30
Dbia PPNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.30
Deug PPNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.32
Drho PPNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.32
Dele PPNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.33
Dfic PPNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.36
Dkik PPNCARCRNHGLKITLKGHKRYCKYRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.49
Dana PPNCARCRNHGLKITLKGHKRYCKYRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.62
Dbip PPNCARCRNHGLKITLKGHKRYCKYRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.63
Dpse PPNCARCRNHGLKITLKGHKRYCKYRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.68
Dper PPNCARCRNHGLKITLKGHKRYCKYRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHMHE 0.68
Dvir PPNCARCRNHGLKITLKGHKRYCKYRYCTCDKCRLTADRQRVMALQTALRRAQAQDEQRSLHMHE 1.00
Dgri PPNCARCRNHGLKITLKGHKRYCKFRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRSLHIHE 1.01
Dwil PPNCARCRNHGLKITLKGHKRYCKYRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRALHIHE 1.01
Dmoj PPNCARCRNHGLKITLKGHKRYCKYRYCTCEKCRLTADRQRVMALQTALRRAQAQDEQRSLHMHE 1.10
B.
Percent of splicing events
100
50
0
F M
Dmel
Dsim
Dbia
Dana
Dpse
Dmoj
Dvir
male-splicing
female-splicing
C.
DSX Occupancy % of Max
Generic Gene Model

Figure 2.4. DSX^F- and DSX^M- occupied regions are not correlated with other transcription factors. Hierarchically clustered heatmap of pairwise similarity metrics between all 255 available ChIP-chip and ChIP-seq experiments and DSX ChIP-seq (highlighted in red) as well as DSX DamID-seq/chip (highlighted in blue). Brighter colors indicate higher similarity (higher fraction of sites with $p < 0.05$); DSX^F and DSX^M (highlighted in red (ChIP-seq) and blue (DamID)) are more similar to each other than they are to any other assayed factor. Self-self comparisons along the diagonal are indicated in gray. Colored blocks along the left side indicate broad clusters. See methods for details. Table S6 contains the source and description of all occupancy data sets tested for correlation with DSX occupancy data.


Fraction of sites with p < 0.05
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Query
Reference

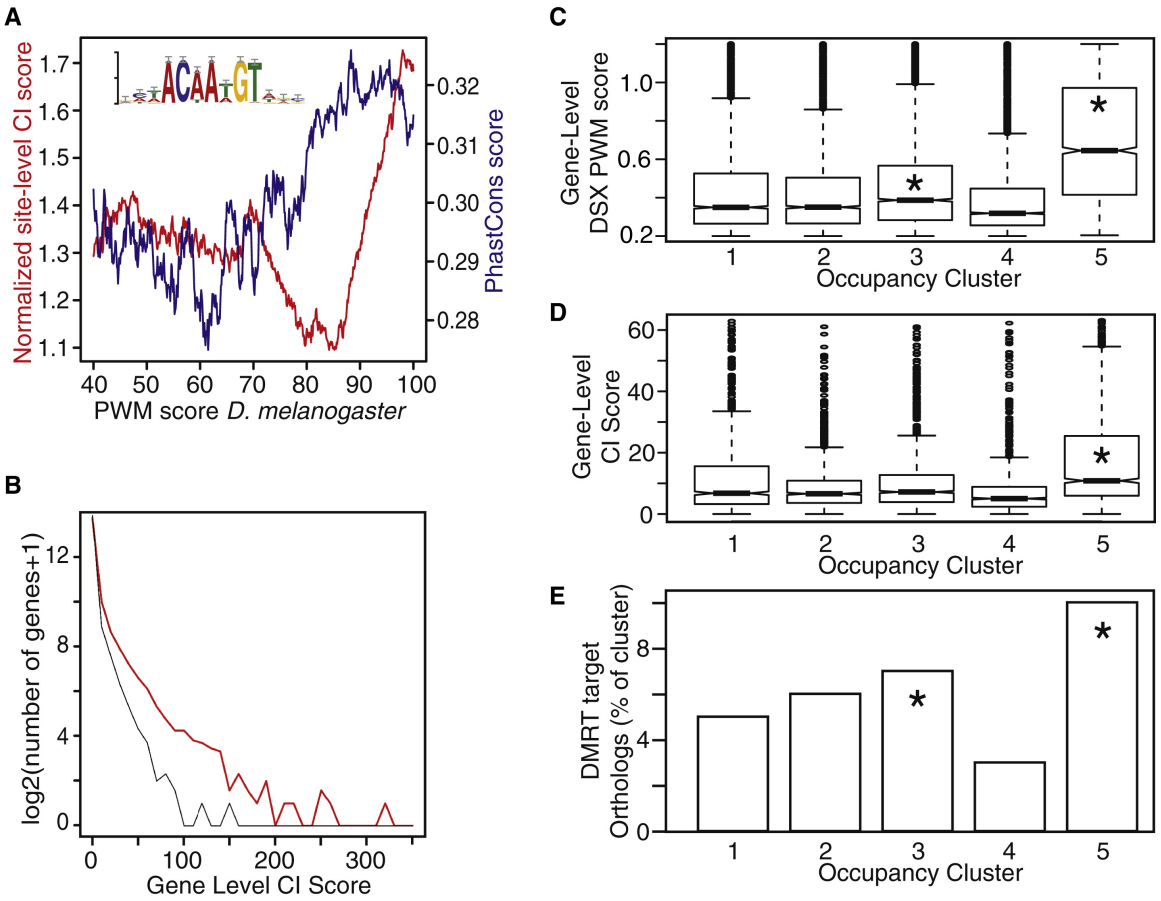
2014; Zhang et al., 2007). In order to analyze the conservation of DSX sites, we performed complementary analyses at both the site- and gene-levels.

The DSX DNA binding domain (DM domain) is highly conserved across animals (Murphy et al., 2007; Yi and Zarkower, 1999), and this domain and sex-specific splicing pattern is conserved across nearly 68 million years of Drosophila evolution (Chen et al., 2014; Figure 2.3). Therefore, we used the same biochemically-defined DSX PWM and 100 position shuffled PWMs as controls to scan the *D. melanogaster* genome and 19 other species in the *Drosophila* genus. We then extracted the *D. melanogaster* sites in the same gene body +1kb range as in the occupancy analysis except that we excluded coding portions of gene bodies to avoid confounding site and codon conservation in the comparative analysis. We assigned each DSX motif in *D. melanogaster* a conservation index (CI) based on the evolutionary distance at which sites could still be identified in the homologous gene using a combination of sequence and distance from the first coding exon (see Materials and Methods, doi: 10.1016/j.devcel. 2014.11.021). We also assigned CI values for the control shuffled motifs using the same method. In order to calculate a CI score for each gene, we summed site-level CIs across a gene body +1kb upstream. We also extracted the well defined and gene length corrected PHylogenetic Analysis with Space/Time models sequences (PhastCons, Siepel et al., 2005) and directly determined the mean PhastCons scores for DSX sites in those segments. Briefly, a high CI or PhastCons score indicates that a site has been conserved, or a new site with similar sequence arose *de novo* at the same relative position. All site- and gene-level PWM and CI scores can be found in Table S2 available at doi: 10.1016/j.devcel.2014.11.021.

As expected, sites that more closely matched the PWM were more likely to have deeper evolutionary conservation (Figure 2.5A). We observed a clear increase in the correlation

Figure 2.5. DSX occupancy and binding-site evolution.

(A) Normalized site-level conservation index scores (taking into account conservation distance and number of species where the site was conserved) plotted against PWM percentile rank score in *D. melanogaster* (red line). Normalized scores for each bin were calculated by subtracting the median site-level CI score for 100 shuffled versions of the DSX motif (see Supplemental Experimental Procedures). PhastCons scores for DSX motifs in *D. melanogaster* (blue line). (B) Histogram of *D. melanogaster* gene-level conservation index scores for DSX (red line) and the median of 100 shuffled DSX motifs (black line). The gene-level CI scores takes into account each motif's conservation index score as well as the number of motifs at the locus. (C-E) We binned genes by occupancy cluster and determined the distribution of gene level DSX position weight matrix (PWM) score (based on the scores and numbers of motifs at a gene) (C), gene level conservation index scores (D), and the percent of genes in each cluster that are orthologs of mouse DMRT1 targets (E). Significant ($p < 0.01$) enrichment using Kruskal-Wallis (B,C) or hypergeometric tests (D) is indicated (asterisks).


A
Normalized site-level CI score
PhastCons score
PWM score D. melanogaster
B
log2(number of genes+1)
Gene Level CI Score
C
Gene-Level DSX PWM score
Occupancy Cluster
D
Gene-Level CI Score
Occupancy Cluster
E
DMRT target Orthologs (% of cluster)
Occupancy Cluster

between PWM score and normalized site-level conservation index, with a prominent "break" in the distribution above the 90th percentile of PWM scores. The PhastCons scores also showed a break in trend but at a lower PWM score rank. This indicates that strong scoring sites show high evolutionary conservation, and are therefore more likely to be functional. Interestingly, in both methods, we observed the poorest normalized conservation in moderately strong *D. melanogaster* sites. The meaning of this dip in the distribution is unclear, but might suggest selection against sites with modest affinity for DSX, which may result in deleterious sex-specific regulation. At the gene-level, DSX CI scores were significantly more conserved (K-S test $p < 2.2e-16$) across evolutionary distance than shuffled PWM CI scores (Figure 2.5B). For this study, we chose to focus our attention on genes with conserved arrays of DSX sites rather than those that were *D. melanogaster* specific due to either species-specific function or chance.

Comparing in vivo occupancy to sequence analysis

We next wanted to evaluate our occupancy clustering analysis based on the sequence-based analysis as a way of focusing on genes that are most likely to be functional DSX targets. As described above, cluster 3 and in particular cluster 5 included those genes with the highest overall occupancy scores (Figure 2.2D). Interestingly, genes in cluster 5 and, to a lesser extent, cluster 3, exhibited significantly greater overall PWM scores than did other clusters (Figure 2.5C). Genes in occupancy cluster 5 also showed significantly higher gene-level CI scores (Figure 2.5D, Table S1 available at doi: 10.1016/j.devcel.2014.11.021), indicating that genes with strong DSX occupancy in *D. melanogaster* had better conservation of DSX binding sites in the *Drosophila* phylogeny. As an estimate as to just how deeply conserved DSX targets might be, we compared our cluster analysis to the list of genes occupied by the mouse DSX homolog,

DMRT1, using whole-genome ChIP (Murphy et al., 2010). Strikingly, orthologs of mouse DMRT1 targets were enriched in DSX occupancy cluster 5, and to a lesser extent cluster 3 (Figure 2.5E). This is somewhat surprising, given the tremendous differences in sexual dimorphism between species. However, perhaps this reflects the fact that DSX/DMRT1 homologs have now been found to control sexual dimorphism across the animal kingdom, acting primarily in the gonads for which sexual dimorphic development is more similar in different species. Overall, our occupancy and sequence analysis are strongly concordant. We therefore focused much of our attention on genes with strong occupancy, strong PWM scores, and strong conservation (Table S1 available at doi: 10.1016/j.devcel.2014.11.021).

Finally, we examined enrichment of gene ontology terms (GO terms) in lists of occupied genes and by occupancy cluster (Table S5 available at doi: 10.1016/j.devcel.2014.11.021) to determine if these gave coherent lists. We found strong enrichment for many different coherent groups of genes in ontologies supporting the idea that DSX controls a wide-range of pathways and functions.

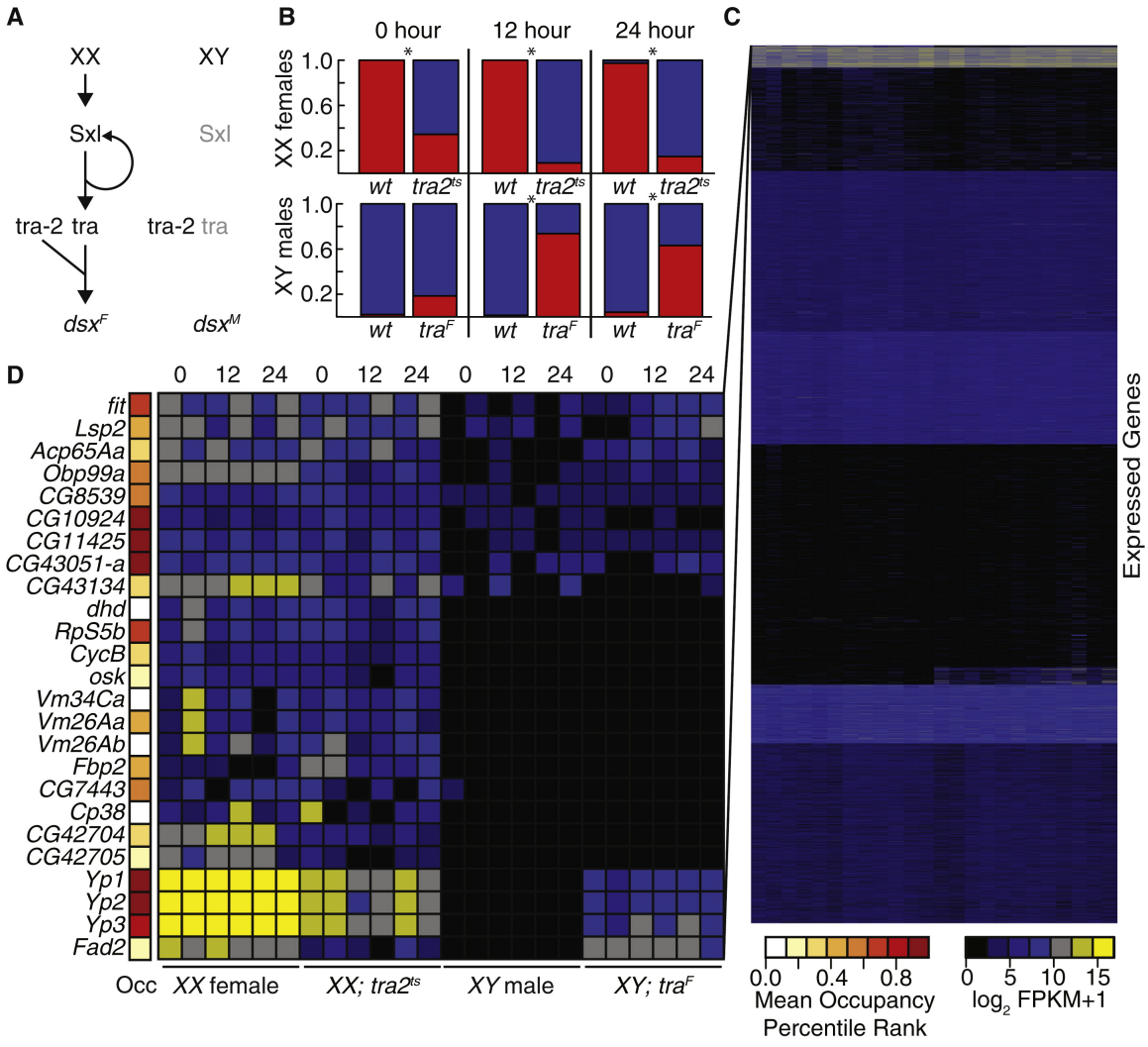
DSX-regulated expression in fat body

Our analysis indicates that many or most DSX target genes exhibit widespread occupancy in the genome that is independent of sex or tissue. However, to control the sex-specific function of many distinct tissues, we expect that DSX^F and DSX^M should have different effects on gene expression with only a subset of targets being relevant in any given sex and tissue. To test this hypothesis, we examined DSX-dependent expression in the adult fat body. We chose fat body because this tissue shows DSX-dependent expression of the *Yp* genes even in adults (Burtis et al., 1991; Coschigano and Wensink, 1993), and because we directly assayed DSX occupancy in this

tissue. To examine DSX-dependent expression, we induced an acute switch in DSX isoform (DSX^F to DSX^M or *vice versa*) using temperature-sensitive alleles of *tra2* or a heat inducible *tra* system (*UAS-tra^F; tub*-GAL4/*tub*-GAL80^{ts}) and performed expression profiling by sequencing (RNA-seq) to test the transcriptional response at 0, 12 and 24 hours following the temperature shift (Figure 2.6, Table S1 available at doi: 10.1016/j.devcel.2014.11.021). We reasoned that switching between DSX isoform states would provide a greater net change in expression level than knocking down DSX function since DSX^M and DSX^F are thought to have opposing roles in target gene regulation (Coschigano and Wensink, 1993). We determined gene-level expression genome-wide and performed k-means clustering to illustrate the overall pattern of expression change in the fat body (Figure 2.6C). 25 genes showed the strongest sex-biased expression (top-level cluster) in these experiments, but only *Yp1*, *Yp2*, *Yp3*, and *Fad2* showed a clear increase in expression correlating with higher DSX^F relative to DSX^M (Figure 2.6D). The response of *Yp* genes to DSX isoforms was as expected based on their known regulation by DSX and confirms that we manipulated known outputs of the DSX pathway (Coschigano and Wensink, 1993; Ronaldson and Bownes, 1995). The *Fad2* locus encodes a female-specific sterol desaturase involved in sex pheromone signaling (Chertemps et al., 2006) that has been reported to be directly regulated by DSX in oenocytes (Shirangi et al., 2009). Our data indicate that DSX also regulates this locus in the fat body, although we observed poor DSX occupancy at the DSX sites (Figure 2.6D, mean fat body occupancy score 25th percentile) raising the possibility of indirect regulation. There were a few genes, such as *CG10924*, *CG11425*, *CG43051-a*, that showed sex-biased expression and strong occupancy. These genes may be regulated by DSX during development or under other environmental conditions but also maintain sex-biased expression using other mechanisms.

Figure 2.6. Tissue-specific DSX function.

(A) Schematic representation of the sex determination cascade in female (XX) and male (XY) flies. Functional mRNAs are indicated in black whereas non-functional mRNAs are indicated in grey. (B) Bar plots of dsx^M (blue) and dsx^F (red) mRNA isoform usage from RNA-seq experiments on adult fat body from controls (wildtype, wt) and experimental flies following temperature shifts. Significant departures ($p<0.001$, Fisher's Exact Test) from control are shown (**). XX; $tra2^{ts}$ flies were morphologically female and fertile while maintained at 18°C indicating that sufficient DSX^F activity existed to support female-specific development and physiology. Similarly, XY; UAS-tra^F/+; *tub*-GAL4/*tub*-$GAL80^{ts}$ flies are phenotypically male and fertile when grown at 18°C. However, we observed some expression of both male and female *dsx* RNA splice forms at the permissive temperature of 18°C in both genotypes. Importantly, upon shifting $tra2^{ts}$ females to 29°C for 12 or 24hrs, *dsx* mRNAs encoding the dsx^M isoform were elevated ~2-fold indicating that we succeeded in manipulating *dsx* isoform-bias in these flies. For the reciprocal acute shift in isoform in UAS-tra^F males, we found that the portion of *dsx* spliced into the female isoform increased ~ 5 fold. (C) Heatmap of gene expression in RNA-seq experiments in adult fat body with sample order fixed. Samples (columns; labeled as in (D)) and genes (rows) are shown. (D) The top-level cluster from (C). Samples times following temperature shift (above) and genotypes (below) are shown in the same order in (C). Genes are listed (left). "Occ" represents the mean occupancy scores from fat body DamID-seq and DamID-array samples. See key for heatmap color-coding.


A
XX
XY
Sxl
Sxl
tra-2 tra
tra-2 tra
dsx^F
dsx^M
B
0 hour
12 hour
24 hour
XX females
XY males
1.0
0.6
0.2
wt
tra2^ts
tra^F
C
Expressed Genes
0.0 0.4 0.8
Mean Occupancy
Percentile Rank
0 5 10 15
log2 FPKM+1
D
0 12 24 0 12 24 0 12 24 0 12 24
fit
Lsp2
Acp65Aa
Obp99a
CG8539
CG10924
CG11425
CG43051-a
CG43134
dhd
RpS5b
CycB
osk
Vm34Ca
Vm26Aa
Vm26Ab
Fbp2
CG7443
Cp38
CG42704
CG42705
Yp1
Yp2
Yp3
Fad2
Occ
XX female
XX; tra2^ts
XY male
XY; tra^F

The most striking finding was that despite the fact that DSX occupies thousands of loci in our adult fat body samples (Table S1 available at doi: 10.1016/j.devcel.2014.11.021), many of which are predicted to be functional based on evolutionary conservation, astonishingly few genes are transcriptionally regulated constitutively by DSX in this adult tissue. This suggests that many genes are poised to respond to DSX but that additional cues (temporal, spatial, nutritional, and/or hormonal) are also required. We conclude that DSX regulatory specificity depends both on where DSX is bound, and the ability of bound DSX to coordinate with other sex-, tissue-, or condition-specific transcription factors or co-factors.

Dose-dependent genetic interactions with dsx

If DSX requires extensive coordination with other inputs as outlined above, then some of the genes occupied by DSX in the fat body should show tissue-specific responses to changes in DSX isoform expression. As a first test of this hypothesis, we conducted an unbiased genetic screen to identify regions of the genome that interact genetically with *dsx* mutations that alter DSX isoform expression (Figure 2.7A-E, see also Table S3 available at doi: 10.1016/j.devcel. 2014.11.021). To compromise *dsx* function, we used the dsx^{D} allele which can only produce DSX^{M}. Consequently, XX; dsx^{D}/+ animals produce both DSX^{F} (from the wildtype allele) and DSX^{M} (from the dsx^{D} allele) which results in an intersexual phenotype similar to that shown by *dsx* loss-of-function (Figure 2.7A-C, Figure 2.8, Nagoshi and Baker, 1990). We tested 101 deletions of the 2nd chromosome (~33% of the genome) to determine which of these regions could modify the XX; dsx^{D}/+ phenotype when heterozygous and analyzed external sex characteristics of the genitalia, abdomen and sex combs (internal structures were not analyzed in

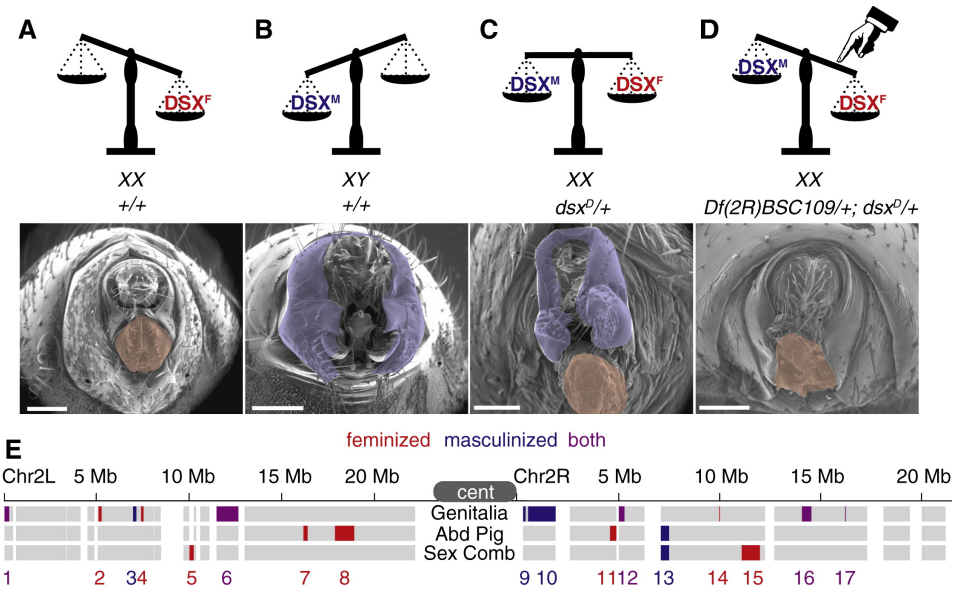


Figure 2.7. Tissue-specific genetic interactions with *dsx*^{D}.

Cartoons (upper) showing DSX isoform in (A) wildtype XX females, (B) wildtype XY males, (C) XX; *dsx*^{D}/+ intersexes, and (D) feminized XX; *dsx*^{D}/+ intersexes due to heterozygousity for *Df(2R)BSC109*/+. Scanning electron micrographs of genitalia (below) showing a major female feature (vaginal plate, red) and a major male feature (genital arch, blue) in false color. Genitalia in XX; *dsx*^{D}/+ flies had female genital structures including a small vaginal plate with fewer teeth than wildtype females (not shown). The vulva failed to close in these intersexes resulting in hemolymph clotting over the vaginal plate. These XX; *dsx*^{D}/+ flies also showed male genital structures including a genital arch that was thin dorsally and spread apart ventrally relative to wildtype males, and with small lateral and posterior lobes. The male clasper teeth were present but reduced in number, and the penis apparatus was usually missing. Scale bar = 100μm. (E) Diagram of the *D. melanogaster* 2nd chromosome with tested regions feminizing (red), masculinizing (blue), feminizing and masculinizing (purple), or having no effect (grey) on intersexual flies shown. Each row outlines the phenotype in genitalia, abdominal pigmentation, and sex combs.

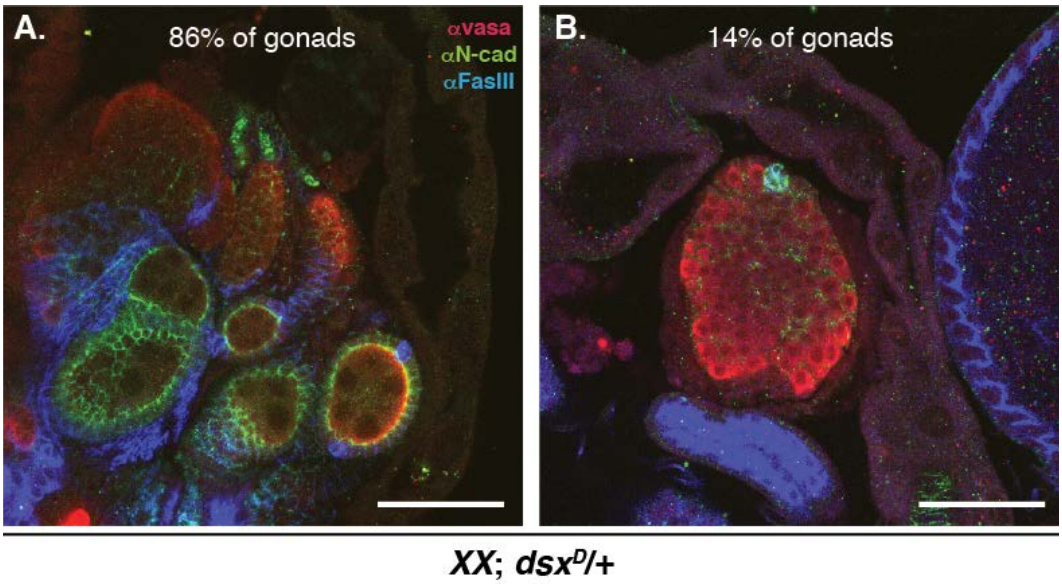


Figure 2.8. XX; *dsx*^{D}/+ gonad phenotypes. Representative images of gonads dissected from XX; *dsx*^{D}/+ adults having either female-like terminal filaments (A) or male-like hubs (B). Terminal filaments and hubs are marked with anti-N-Cad (green), hubs are marked with anti-FasIII (blue) and anti-N-cad (green), and germ cells with anti-Vasa (red). Scale bar = 50μm.

this initial screen) to determine if genetic interactions where tissue-specific (Table S3 available at doi: 10.1016/j.devcel.2014.11.021).

Indeed, a number of 2nd chromosome deletions exhibited the tissue-specific genetic interactions with *dsx* (Figure 2.7E). For example, in XX; *Df(2R)BSC109/+; dsx*^{D}/+ flies, all male-like genital structures were missing and female genital structures were more pronounced, including a larger, fully-closed vaginal plate replete with teeth (Figure 2.7D); however, there were no changes in sex comb morphology, tergite number, or abdominal pigmentation. Thus, these data suggest that a gene (or genes) in the *Df(2R)BSC109* region is required, in conjunction with *dsx*, for male development of the genital disc but not in other tested tissues. Of the 101 deletions tested, we identified 19 *Dfs*, defining 17 unique genomic intervals that modified the *dsx*^{D}/+ phenotype in one of the three tissues examined (Figure 2.7E). Strikingly, only a single region affected sex differentiation in more than one tissue, and this *Df* removed *intersex*, which encodes a protein that binds DSX^{F} and is thought to be important for all aspects of DSX^{F} function (Garrett-Engele et al., 2002). The remaining 16 interacting regions modified the *dsx*^{D}/+ phenotype in only a single tissue. We conclude that genes interacting with *dsx* do so in a highly tissue-specific manner. This supports the idea that, although DSX binds many of the same loci in different tissues, it regulates distinct downstream targets to control sexual differentiation in these different tissues.

Tissue-specific effects of predicted DSX targets

Since our deletion screen indicated that most loci interacting with *dsx* do so in a highly tissue-specific manner, we wanted to determine if this was true for DSX target genes. To do this, we selected 60 predicted DSX targets from our analysis and examined their loss of function

phenotype using RNAi. These were selected primarily due to high occupancy, PWM scores, and conservation. We also biased the set to named genes with existing alleles, and selected some genes based on other criteria such as localization to a *Df* region interacting with dsx^D. See Table S4 available at doi: 10.1016/j.devcel.2014.11.021 for a full list of locus characteristics contributing to selection including occupancy cluster, gene-level PWM score, gene-level CI score, and orthology to mouse DMRT1 targets. To restrict our analysis to tissues that are regulated by *dsx*, we expressed these UAS-RNAi constructs using *dsx*-GAL4 (Rideout et al., 2010; Robinett et al., 2010) and examined the sexual morphology of 16 sexually dimorphic structures (see Table 2.4). The genes we selected for analysis were not random since they depended on the availability of existing reagents and other criteria. Thus, this was not a random screen but is still informative as to the nature of loss of function phenotypes associated with predicted DSX targets.

As in the dsx^D interaction screen, we observed striking tissue-specific loss of function phenotypes in sexually dimorphic tissues (Table 2.4). For example, *thickveins* (*tkv*) knockdown resulted in an increase in male-like abdominal pigmentation in females similar to those observed in dsx^{RNAi} females (Figure 2.9A), but showed no effect in any other tissue in either sex. In gonads, *abd-A* knockdown females exhibited disorganized ovaries that failed to attach to the rudimentary genital tract but no testis phenotype, while *bunched* (*bun*) knockdown males exhibited a bulbous testis but no ovary phenotype (Figure 2.9B). Another clear tissue-specific sex transformation was observed in *neuralized* (*neur*) knockdown females (Figure 2.9C), which showed the male-specific large central bristle below the normal-looking female leg bristles. In males, knockdown of *chameau* (*chm*) resulted in pointed sex comb teeth, as observed in females, but the sex combs showed male thickness, rotation, and pigmentation (Figure 2.9C) indicating

Figure 2.9. Tissue-specific functions of DSX target genes.

(A) Abdominal pigmentation phenotypes of wildtype female, wildtype male, *dsx* knockdown male, and *thickveins* (*tkv*) knockdown female (left to right). (B) Gross anatomy of gonads from wildtype female, *abdominal A* (*abd-A*) knockdown female, wildtype male, and *bunched* (bun) knockdown male. Terminal filaments and hubs are marked with anti-N-Cad (green), somatic gonadal cells with anti-traffic jam (TJ, blue), and germ cells with anti-Vasa (red). Scale bar = 50μm. (C) First leg tarsal segments from wildtype female, wildtype male, *neuralizer* (*neur*) knockdown female, *dsx* knockdown male, and *chameau* (*chm*) knockdown male (left to right). The male-specific central bristle is shown (arrowhead). (D) Scanning electron micrographs of genitalia from wildtype female, *abdominal-A* (*abd-A*) knockdown female, wildtype male, and *bunched* (*bun*) knockdown male. Scale bar = 50μm.

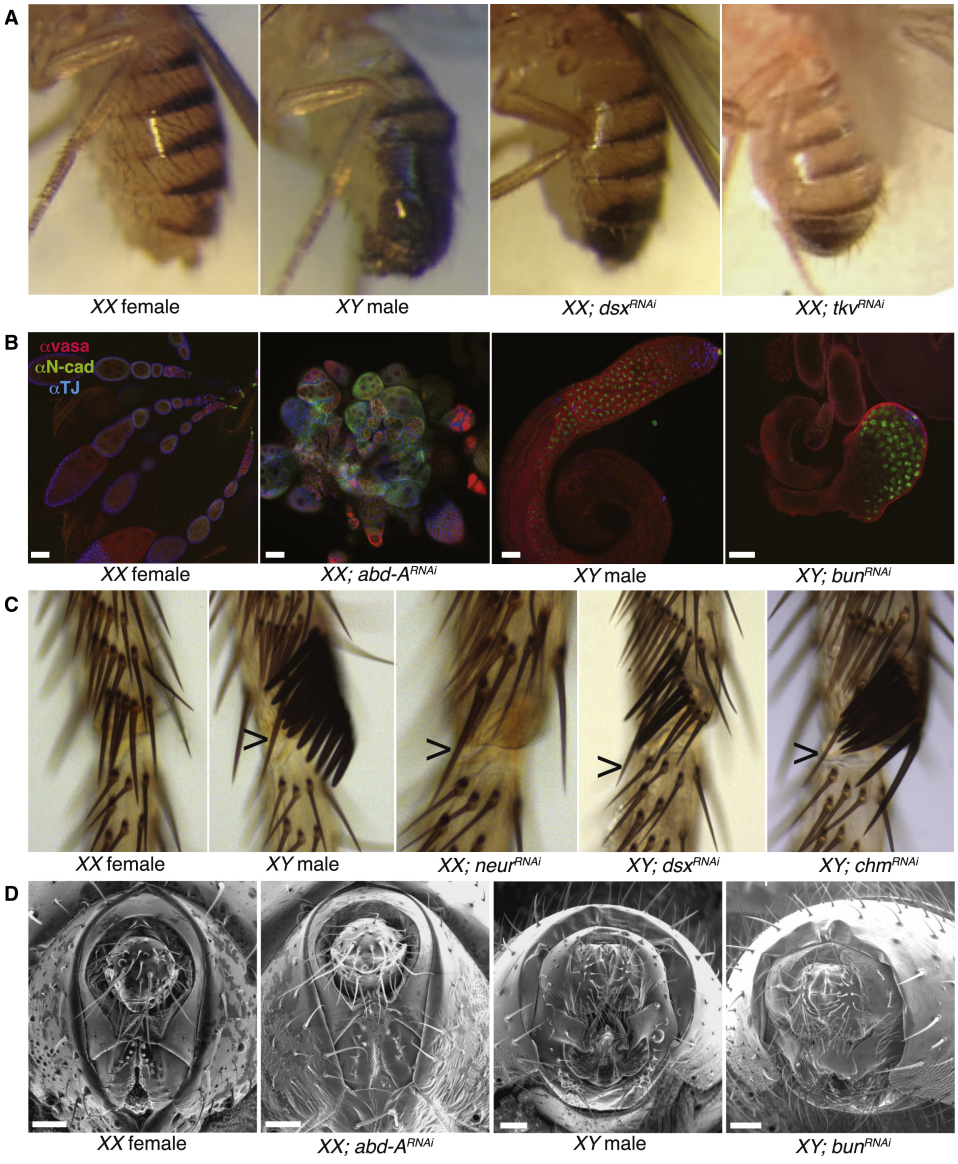
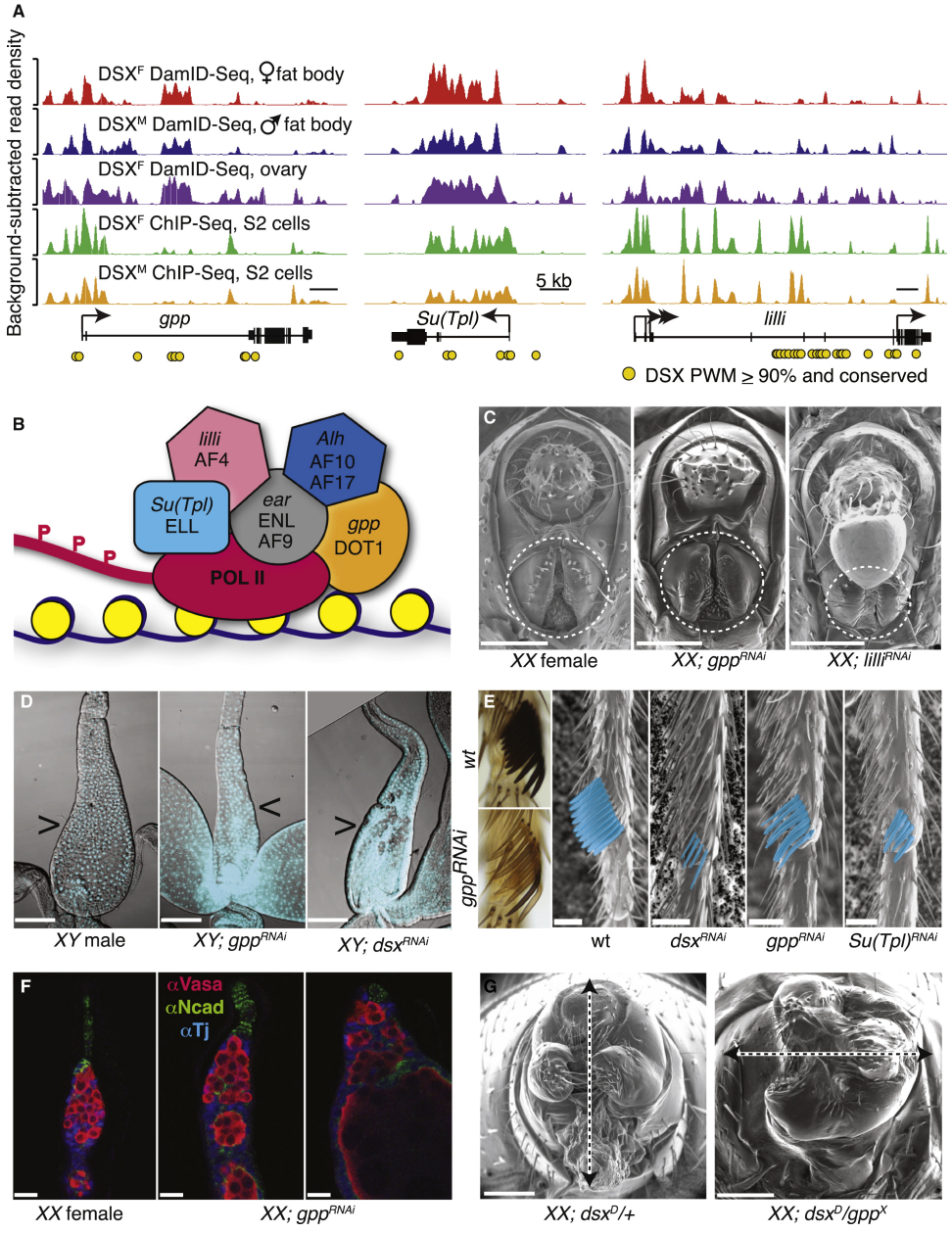

A
XX female
XY male
XX; dsx^RNAi
XX; tkv^RNAi
B
αvasa
αN-cad
αTJ
XX female
XX; abd-A^RNAi
XY male
XY; bun^RNAi
C
XX female
XY male
XX; neur^RNAi
XY; dsx^RNAi
XY; chm^RNAi
D
XX female
XX; abd-A^RNAi
XY male
XY; bun^RNAi

Figure 2.10. Function of DOT1 in sex differentiation.

(A) Scaled read density plots from five replicated DSX occupancy experiments (as labeled) along the genome for *grappa* (*gpp*), *Suppressor of Triplolethal* (*Su(Tpl)*), and *lilliputian* (*lilli*) loci. FlyBase gene models are shown (as in Figure 1) as are positions of high scoring (≥90th percentile) DSX binding sites some of which were conserved in at least one other Drosophila species (yellow circles). (B) Cartoon showing DOT1 and associated yeast proteins (capital letters) loaded onto elongating RNA polymerase (see Discussion). *Drosophila* genes encoding orthologs are in italics. (C) Scanning electron micrographs of wildtype female, *gpp* knockdown female, and *lilli* knockdown female (left to right) genitalia. The vaginal plate and teeth are highlighted (dotted circles). Scale bar = 100μm. (D) Light microscopy images wildtype, *gpp* knockdown, and *dsx* knockdown (left to right) male ejaculatory ducts (arrowheads) stained with DAPI (light blue). Scale bar = 100μm. (E) Male first leg tarsal segments. Light micrograph of wildtype (left top) and *gpp* knockdown male (left bottom) sex combs. Scanning electron micrographs (last four panels) of wildtype and *dsx*, *gpp*, and *Su(Tpl)* (left to right) male sex combs (teeth false colored). (F) Confocal micrographs of wildtype and two examples of *gpp* knockdown female germline stem cell niches. Terminal filaments (anti-N-Cad; green), somatic cells (anti-Traffic Jam; blue), and germ cells (anti-Vasa; red). Scale bar = 10μm. G) XX; dsx^{D}/+ intersex control or heterozygous for *gpp* (left to right). The axis of incomplete rotation in dsx^{D}/ gpp^{X} male genitalia is shown (dotted). Scale bar = 100μm.


A
Background-subtracted read density
DSX^F DamID-Seq, ♀ fat body
DSX^M DamID-Seq, ♂ fat body
DSX^F DamID-Seq, ovary
DSX^F ChIP-Seq, S2 cells
DSX^M ChIP-Seq, S2 cells
5 kb
gpp
Su(Tpl)
lilli
DSX PWM ≥ 90% and conserved
B
lilli AF4
Alh AF10 AF17
Su(Tpl) ELL
ear ENL AF9
gpp DOT1
POL II
P P P
C
XX female
XX; gpp^RNAi
XX; lilli^RNAi
D
XY male
XY; gpp^RNAi
XY; dsx^RNAi
E
wt
gpp^RNAi
wt
dsx^RNAi
gpp^RNAi
Su(Tpl)^RNAi
F
αVasa
αNcad
αTj
XX female
XX; gpp^RNAi
G
XX; dsx^D/+
XX; dsx^D/gpp^X

that multiple pathways regulate the wildtype male sex comb phenotype. In addition to the ovary phenotype, *abd-A* knockdown females displayed genitalia defects with recessed vaginal plates with reduced numbers of vaginal teeth. Similarly, *bunched* (*bun*) knockdown males were missing the penis apparatus and the majority of clasper teeth (Figure 2.9D) in addition to the testis phenotype. Finally, we also observed cases where knockdown of the same target gene resulted in defects in female differentiation in one tissue and male differentiation in another tissue. For example, *longitudinals lacking* (*lola*) knockdown females were almost entirely lacking external genitalia, while males had wide, bulbous testes (Table 2.4). Thus, the RNAi results demonstrate that genes that are bound by DSX in multiple tissues can have striking tissue-specific effects.

DOT1 complex

We might expect that multiple genes in a common pathway or that encode members of the same protein complex might be co-regulated by DSX and would exhibit similar signatures as DSX targets with related loss of function phenotypes. Indeed, we found that many of the genes encoding the Disruptor Of Telomeric silencing-1 (DOT1) complex(es) (Biswas et al., 2011; Bitoun et al., 2007; Mueller et al., 2007; Okada et al., 2005; Zeisig et al., 2005) are predicted DSX targets (Figure 2.10A). The DOT1 complex acts as a positive transcriptional regulator through methylation of histone H3 at lysine 79 along gene bodies in many studied organisms (Feng et al., 2002; Nguyen and Zhang, 2011). In *Drosophila*, DOT1 is encoded by *gpp* (Shanower et al., 2005) and members of the complexes are encoded by *lilliputian* (*lilli*), *ENL/ AF9-related* (*ear*), *Alhambra* (*Alh*) and *Supressor of triplolethal* (*Su(Tpl)*) (Figure 2.10B). We observed strong DSX occupancy at *gpp, Su(Tpl)*, *lilli,* and *Alh*, but not *ear* (Figure 2.10A, Table

S1 available at doi: 10.1016/j.devcel.2014.11.021). DSX^F occupancy at *Alh* and *lilli* was also previously reported (Luo et al., 2011). Furthermore, the DSX binding sites for these genes were well conserved in the *Drosophila* phylogeny, and the mouse orthologs of *Su(Tpl)* and *lilli* are occupied by mouse DMRT1 (Murphy et al., 2010), suggesting that the DOT1-containing complexes are evolutionarily conserved targets of DMRTs. Given that these proteins function together in a variety of complexes, mutations should result in similar sex-transformation phenotypes.

To determine if DOT1 functions in *dsx*-expressing cells, we knocked down the DOT1 complex-encoding genes using *UAS-RNAi* constructs driven by *dsx*-GAL4. Overall, we observed sex- and tissue-specific phenotypes among genes encoding members of the complex except *Alh,* which did not result in an overt phenotype. Externally, we found that the female genitalia of *gpp, Su(Tpl)*, and *lilli* knockdown flies had reduced or no vaginal teeth on the otherwise normal genitalia (Figure 2.10C, Table 2.4), while the male genitalia were grossly defective, with missing lateral lobes and claspers, and a missing penis apparatus. Additionally, the male (but not the female) genital disc rotates along the anterior/posterior axis during development (Adam et al., 2003), and we observed partially rotated genitalia in DOT1 knockdown males (Table 2.4). Internally, the parovaria and spermathecae along the internal female reproductive tract were missing from *gpp* knockdowns. In males, *gpp* knockdown resulted in a narrow and thin ejaculatory duct similar to those observed following *dsx* knockdown (Figure 2.10D). The female and male internal reproductive structures derive from different segments of the genital disc (reviewed by Estrada et al., 2003), suggesting that *gpp* has sex- and segment-specific roles in both internal and external genital development. Males also showed a specific defect in the sex combs. Knockdown of *gpp*, *Su(Tpl), lilli*, or *ear* in males

resulted in decreased pigmentation of the sex combs (Figure 2.10E, Table 2.4). Additionally, *gpp* and *Su(Tpl)* knockdown also resulted in reduced numbers of sex comb bristles with a thinner and more pointed appearance, suggesting that they were partially feminized (Figure 2.10E). Interestingly, not all aspects of male pigmentation were affected, since male-like pigmentation of abdominal segments was normal (Table 2.4). Lastly within the gonads, *gpp* knockdown in the ovary altered the morphology of the niche region, where we observed collapsed terminal filaments and excessive numbers of early stage germ cells (Figure 2.10F), while the male niche was unaffected. This suggests that *gpp* is required for female niche development. In summary, the collections of specific defects observed in these RNAi experiments indicate that members of the DOT1 complex(es) have similar sex- and tissue-specific functions in *dsx*-expressing cells.

The fact that the DOT1 complex members show sex- and tissue-specific phenotypes when knocked down is consistent with them being DSX target genes; however, an alternative interpretation is that the DOT1 complex(es), which function broadly at active genes, is important for proper gene regulation in general, and that by knocking down the function of this complex in *dsx*-expressing tissues we are simply interfering with the sex-specific development of these tissues in a manner unrelated to *dsx* function. To address this, we examined the genetic interaction between *dsx* and mutant alleles of genes encoding DOT1 complex members. If the DOT1-containing complex(es) acts together with *dsx*, we might expect a dosage sensitive interaction between *dsx* and genes encoding complex members, whereas if the DOT1 complex(es) acts independently of *dsx* we should not. To test this we again used the dsx^D background that creates an intersexual phenotype in *XX;* dsx^D/+ animals. When we introduced heterozygosity for gpp^X in the dsx^D background, we observed reduced male genitalia structures and defective genitalia similar to the RNAi knockdown phenotypes for *gpp* and *lilli* (Figure

2.10G, Table 2.4). In addition, XX; dsx^D/+ animals show a phenotype in the gonads where either male structures (the hub) or female structures (terminal filaments) are present (Figure 2.8). We observed an increased rate of hub formation in dsx^D gonads after reducing the dose of either *gpp* (gpp^X, 62% hub, n=37) or *lilli* ($lilli^{A17\text{-}2}$, 39% hub, n=36) relative to those dsx^D/+ gonads (14% hub, n=106) suggesting that these genes play a positive role in terminal filament development or a negative role in hub development. Thus, like the RNAi experiments, the genetic interactions suggest that the DOT1 complex(es) are required for female niche development. In conjunction with the genomic data, the RNAi and dsx^D interactions strongly suggests that the genes encoding the DOT1 complex(es) are directly and coordinately regulated by DSX in specific tissues.

Discussion

Finding functional DSX targets

Determining which genes are directly regulated by any transcription factor is complicated by the fact that they recognize short sequences in the genome that can arise by chance. The use of multiple genome-wide techniques helps winnow potential targets. To understand how DSX contributes to sex- and tissue-specific development and to catalog DSX targets genes in tissues that express *dsx*, we undertook a series of genome-wide experiments and analyses to determine: where DSX is bound in different cell types, which sites are evolutionarily conserved in 20 sequenced *Drosophila* species and in the mouse, the relationship between site strength and occupancy, and which genes respond to acute changes to DSX^F/DSX^M isoform abundance. We then performed RNAi knockdown and dosage-sensitive genetic interaction tests of candidate targets and found that they resulted in striking tissue-specific transformations of subsets of sexually dimorphic structures. Among predicted targets, we found enrichment for GO terms

involved in regulation of transcription (adjusted p-value=1.85E-7) and signaling (adjusted p-value=1.44E-53), suggesting that DSX regulates gene expression of terminal differentiation factors by direct and indirect mechanisms. For example, candidate DSX transcription factor targets include *apterous* (*ap*), *brother of odd with entrails limited* (*bowl*), *Antennapedia* (*Antp*), *chronologically inappropriate morphogenesis* (*chinmo*), *taiman* (*tai*), *mastermind* (*mam*), and *Zn finger homeodomain 1* (*zfh1*) (Table 2.4). Candidate DSX signaling targets include *dachsous* (*ds*), *frizzled* (*fz*), *frizzled 2* (*fz2*), *decapentaplegic* (*dpp*), *Leucokinin receptor* (*Lkr*), *Ecdysone Receptor* (*EcR*), and *Dopamine 1-like receptor 2* (*DopR2*).

This rich set of DSX target genes will be useful for ultimate *in toto* understanding of the sex differentiation network in the powerful Drosophila system. The enrichment for orthologs of genes bound by mouse DMRT1 among the DSX targets with highest occupancy strongly suggests that some of this network will be conserved in mammals. We find that the integration of multiple data sets such as transcription factor occupancy and binding site conservation is a powerful approach that can be applied to any transcription factor and is especially attractive in Drosophila where 20 species have been sequenced representing 68 million years of evolution.

The logic of DSX regulation

A large number of target genes might suggest that DSX acts as a "micro-manager" of sexual development, regulating the expression of many or most terminal sex-differentiation genes, such as *Yp1*, *Yp2*, (Burtis et al., 1991) and *Yp3* (this study). However, our unbiased screen to identify genes interacting with *dsx*, using deletions in a dsx^D genetic background, predicts a smaller number (≈50) of "major effect" loci acting in the *dsx* pathway. How can we reconcile the disparity between the large numbers of genes exhibiting DSX-occupied, evolutionarily

conserved DSX sites and the many fewer predicted to have “major” effects? Perhaps DSX delegates regulatory function to major pattern formation pathways that lead to sex-specific development of organ systems such as the gonad, reproductive tract, and CNS. This might account for the large number of transcriptional regulators that show DSX regulated sex-biased expression (Chatterjee et al., 2011) which are predicted DSX targets (this study). In addition, many genes regulated by DSX might provide more subtle, but still evolutionarily significant, "minor" polygenic effects on development or physiology. DSX regulation of these minor, polygenic effect loci could help explain the effects of genetic background on sex-related phenotypes. These major and minor effect genes would both be strongly selected for in the course of evolution.

Types of DSX targets

The types of genes that are predicted by our analyses illustrate how DSX is able to exhibit such powerful effects on developmental pathways. One group of predicted target genes are genes involved in short-range (e.g. WNT, EGF, and DPP), and long-range (e.g. Insulin and Ecdysone) signaling. Thus, DSX expression in a small subset of cells could have far-reaching effects on the development of surrounding cells and beyond. Indeed, genital disc (Ahmad and Baker, 2002; Gorfinkiel et al., 2003; Keisman et al., 2001) and gonad (DeFalco et al., 2008; Oliver et al., 1993; Wawersik et al., 2005) development studies have demonstrated that DSX modulates short-range signaling pathways to result in non-autonomous phenotypes. Finally, coordinated production of the Yolk Proteins in the fat body requires hormonal communication between the fat body and the ovaries in addition to DSX (Bownes et al., 1996). Titers of the steroid ecdysone are highly female-biased and germline-dependent in adult *Drosophila* (Parisi et

al., 2010) suggesting that sex determination genes regulate the hormonal inputs. Our findings may help close this physiological loop by suggesting that DSX is a direct transcriptional regulator of hormonal signaling pathways.

One advantage of having DSX act on signaling pathways is that this provides a mechanism for cells to "agree" on which sex-specific developmental path to take, allowing the sexual decision of a tissue to be reinforced and maintained. Such mechanisms are common in sex determination, such as in the mammalian gonad where one of the most immediate effects of sex determination is to activate sex-specific signaling pathways (reviewed by Eggers and Sinclair, 2012) which is thought to reinforce and provide uniformity in the sex determination process. Interestingly, the WNT and FGF pathways are also involved in sex-specific gonad and reproductive tract development in *D. melanogaster* (Ahmad and Baker, 2002; DeFalco et al., 2008). While *Drosophila* and mammalian sex determination is often thought of as being very different, the emergence of these overlapping modules of gene interactions suggests significant commonalities.

Another major class of potential DSX targets we identified encode transcriptional regulators. Previous genome-wide and case-by-case studies have shown sex-specific expression patterns of transcriptional regulators in dimorphic tissues such as genital disc, leg discs, and abdominal histoblasts (Barmina et al., 2005; Chatterjee et al., 2011; Williams et al., 2008). Transcription factors are also among the potential DSX^F target genes (Luo et al., 2011). By influencing transcription factors, the action of DSX can be amplified to induce far-reaching developmental programs. We suggest that DSX is not generally a "micro-manager" of developmental programs, but instead "delegates" to activate pathways that proceed largely without further input by DSX. There is clear evidence for this mode of action. In the absence of

dsx function, both male and female reproductive structures are found in adults. This is opposed to the absence of all sexual structures that we would expect if DSX were a "micro-manager" of their development. Further, within the gonad, in the absence of *dsx* function, components of either the male or female gonad stem cell niches, the hub and terminal filaments, still form, but do so stochastically in XX and XY flies (N. Camara, CW and MVD, unpublished). We observed a similar all-or-nothing niche phenotype in XX; dsx^{D}/+ flies. Again, we would expect these structures to be absent if *dsx* was required for their formation. These data are consistent with DSX being a regulator of other regulators that control female- and male-specific development.

We also found co-factor complexes and epigenetic transcriptional regulators among the potential DSX target genes suggesting potential "fine tuning" and/or "memory function" for DSX. For example, the DOT1 epigenetic machine mediates mono-, di-, and tri-methylation of H3K79 in multiple species including *D. melanogaster* (Feng et al., 2002; Shanower et al., 2005). H3K79me is enriched along the gene body of highly expressed genes (Steger et al., 2008) although in Drosophila, DOT1 (*gpp*) may have roles in both activating and repressing transcription in a domain- or gene-specific manner (Shanower et al., 2005). While DSX control of DOT1 epigenetic action could act as cellular memory systems and/or generally boost expression of a large set of genes, it may also be functioning to regulate gene expression of only one or a few genes that contribute to a sexually dimorphic phenotype. For example, the Dot1L target gene dystrophin appears to be the primary gene contributing to the cardiac defects in Dot1L knockout mouse cardiomyocytes as re-introduced dystrophin rescued the cardiac Dot1L phenotypes (Nguyen et al., 2011).

In addition to general transcriptional co-factors and epigenetic modifiers, feedback systems and cross-regulation also affect the output levels and stability of genetic pathways. It is

striking that one class of genes predicted to be DSX targets are members of the sex determination hierarchy itself. Both *dsx* and *fru* have evolutionarily-conserved DSX (this study) and FRU (Neville et al., 2014) binding sites. Further, *Sxl*, which regulates *tra* (reviewed by Camara et al., 2008), has conserved DSX binding sites. It is possible that DSX regulates *Sxl* expression in some tissues (although none are known) and this could ultimately influence *dsx* alternative splicing via TRA. There is precedent for feedback in the sex determination "hierarchy" as TRA is a feedback regulator of *Sxl* (Siera and Cline, 2008). Similarly, predicted targets of DSX such as *Scr*, *Abd-B*, and others (this study) encode transcription factors that are known to regulate *dsx* expression (Chatterjee et al., 2011; Devi and Shyamala, 2013; Tanaka et al., 2011; Wang and Yoder, 2012). Recent demonstration of sex determination modulation by micro-RNAs (Weng et al., 2013) suggests that we are quite far from a clear understanding of even the basic framework for sex determination and differentiation. We suggest that sex differentiation occurs via a set of context-dependent networks -- replete with rich auto-regulation, cross-regulation, and feedback -- not a hierarchy.

In summary, the wiring diagram surrounding *dsx* may be quite complex as DSX directly or indirectly regulates a broad set of transcription factor encoding genes including some of those that regulate *dsx* expression. If DSX is both regulated by, and a regulator of, a broad array of transcription factors that are widely deployed during development, then inappropriate expression of DSX could be deleterious. Indeed, ectopic expression of *dsx* results in widespread changes in morphology and in lethality (Jursnich and Burtis, 1993), suggesting that *dsx* must be tightly regulated. The *dsx* gene is expressed in a highly tissue-specific manner (Hempel and Oliver, 2007; Lee et al., 2002; Rideout et al., 2010; Robinett et al., 2010) consistent with the idea that only those tissues that have sex-specific developmental programs express *dsx*. Understanding

the logic by which DSX acts to control distinct dimorphic developmental outcomes in different tissues in the context of multiple highly integrated networks a key question in *Drosophila* sex determination.

Specificity of DSX action

In a multiple network model, DSX must act in different tissues to regulate the diverse sexual processes for which it is required. The sex-specific developmental programs of the gonadal mesoderm, the leg or genital imaginal discs, the fat body, and the nervous system are all likely to be highly divergent, yet all depend on DSX. How?

A priori, DSX could regulate different target genes in different tissues and times using any combination of a host of locus characteristics. First, despite the presence of a common DNA binding motif, the DSX^F and DSX^M isoforms could have distinct binding sites. While we do find genes with sex-specific occupancy patterns, this model is not well supported. Second, DSX could bind different genes in different tissues due to co-factor requirements or site availability in chromatin. While we found examples of tissue-biased DSX occupancy (e.g. *Yp1*, *Yp2*, and occupancy clusters 1 and 2), the vast majority of candidate targets were occupied in all tested tissues and cells. Third, DSX could always bind a similar target gene battery, but the ability to regulate gene expression would depend on the combinatorial activity of other gene-specific transcription factors. Our work provides strong evidence for the last model, although we do not rule out the first two. There is also support for this combinatorial model in the literature. The *bric-a-brac1* (*bab1*) locus is regulated by an enhancer that bears both DSX and homeobox (HOX) protein binding sites to control sex-specific expression along the anterior/posterior axis (Williams et al., 2008). And, as we discussed earlier, *Yp1* and *Yp2* are regulated by DSX (Burtis

et al., 1991) in conjunction with Ecdysone signaling via the same short intergenic region between these divergently transcribed genes (Bownes et al., 1996).

Our data shows that DSX function is primarily regulated by mechanisms other than where it is recruited to DNA for the majority of predicted targets. We find that DSX is bound at largely overlapping sets of genes, regardless of the tissue being analyzed or the DSX isoform. This is true even in samples from S2 cells that do not express endogenous *dsx* above background levels thus providing an exogenous chromatin context for DSX binding. Further, although we observed DSX binding at thousands of genes in the fat body, and we found that many of these DSX sites are conserved in the Drosophila phylogeny, we observed only a handful of DSX target genes with robust expression changes in the adult fat body when the DSX isoform was acutely switched from male to female, or *vice versa*. Thus, only a few genes were functionally regulated by DSX, despite DSX occupancy, even though both measurements were made at a matched developmental time, place, and experimental condition. We conclude that, in the contexts we examined, DSX binding at a gene confers the possibility of sex-specific regulation; however, functional regulation of a target requires other inputs such as tissue, temporal, spatial, and/or hormonal factors. We did not examine occupancy and expression changes throughout developmental time, but our results lead to the prediction that there are likely to be distinct transcriptional responses to DSX among tissues throughout development even though the occupancy patterns are likely to be largely similar.

Even providing sexual information is more complicated than we anticipated. For example, in females *gpp* is required for development of vaginal teeth, while in males *gpp* is required for development of sex-combs. This suggests that the sexual directionality of DSX^F and DSX^M regulation of *gpp* depends on tissue-specific co-factors. Clearly, for the majority of

targets examined in this work, the DSX influence on gene expression depends on mechanisms independent of where DSX binds in the genome. We do not suggest that this pattern of occupancy and regulation is the only way to achieve tissue-specificity of targets, but rather that this pattern is employed in a large number of DSX targets.

DSX and DOT Complexes

We identified far more candidate DSX target genes than we could address within the scope of a single study. We chose to investigate the DOT1 encoding *gpp* gene because it is a member of one or more complexes. As mutations in any of the members is expected to impact complex(es) function, we expected similar sex transformation phenotypes by manipulating expression in RNAi experiments or by genetic interactions with dsx^{D}/+. We observed strong DSX occupancy and conserved binding sites at *gpp*, *Su(Tpl)* and *lilli*. The similar phenotypes we observed in flies expressing RNAi constructs to knockdown *gpp* (DOT1), *Su(Tpl)* (ELL) and *lilli* (AF4) suggests that these proteins have a similar functions in *Drosophila*, which is in turn consistent with membership in common complex(es) regulated by DSX. The *gpp* RNAi and *dsx* interaction phenotypes were more extreme, which is expected if DOT1 is a member of multiple complexes.

The *gpp* locus was identified in a genetic screen for pairing-dependent Polycomb silencing, leading to an interpretation of *gpp* phenotypes in terms of segment identity (Shanower et al., 2005). Some defects in *gpp* mutants, such as transformation of arista to legs, are clear examples of defects in segmental identity, but we suggest that other phenotypes relate to sex differentiation defects. For example, *gpp* mutants have reduced pigmentation of the male sex combs bristles as seen in our experiments driving *gpp* RNAi constructs in *dsx*-expressing cells.

We did not observe abdominal sex differentiation phenotypes due to *gpp* knockdown in our study, but a vestigial abdominal segment 7 (A7) has been observed in *gpp* mutant males. This is significant as A7 is present in females and is not visible in males. Additionally, *gpp* males show reduced pigmentation of A5 and A6, which is consistent with either segment identity change or sex differentiation. Mutations in *gpp* also showed genetic interactions with *Abd-B* resulting in a genital disc rotation defect in males (Shanower et al., 2005). Interestingly, we also observed a genital disc rotation phenotype, in dose-dependent genetic interaction between *gpp* and dsx^D. We suggest that the role of *gpp* in segment identity includes sexual dimorphism.

In over 100 years of studying sex determination and differentiation, only a few key genes have been identified. Our work provides a rich set of DSX target genes for future studies and outlines the mechanisms of DSX action in broad strokes. DSX sex-specific isoforms often bind the same genes, where context-specific factors determine the consequences of that binding. These complex, context-dependent patterns mean that DSX^F can act as a positive regulator of a target gene in one tissue, and DSX^M can act a positive regulator of the same locus in another. DSX acts by a combination of delegating control to transcription factor target genes and by directly micromanaging terminal differentiation genes in a tightly integrated dance of regulatory inputs. While we may still have decades of research on the roles of DMRT genes in sex determination and differentiation, we now have a comprehensive target gene resource to guide that effort.

CHAPTER 3: RNA-SEQ TO IDENTIFY GENES CONSTITUTIVELY REGULATED BY DSX IN ADULT FAT BODY AND GONADS

Summary

In the past decade, few comprehensive approaches have been used to identify genes that are sex-differentially expressed and regulated by the sex-determination hierarchy in adults. Studies primarily using microarrays and serial analysis of gene expression (SAGE) techniques have examined mRNA profiles of tissues (heads and CNS) from flies mutant for *tra*, *dsx* and *fru* at different developmental stages (pupae and adults). These techniques were unable to discover genes deployed through the action of *dsx* in adults. In this chapter, we use conditional mutants of the *Transformer* gene involved in sex determination, to temporally control which DSX isoform will be produced and examine mRNA profiles of gonads from adult females that have undergone the switch from DSX^F to DSX^M (and vice versa) to identify those genes that are being regulated through the action of sex-specific isoforms of DSX.

To gain a comprehensive understanding of biologically significant targets of DSX, we combined the dataset produced from our occupancy approaches (Chapter 2) with our expression profiling dataset from adult flies that have undergone a switch in DSX isoform for an acute period of time. We find that although thousands of genes are occupied by DSX, only a subset of genes are transcriptionally regulated by DSX. Thus, we find that genes are poised to respond to DSX, but that additional cues are also require. We conclude that DSX regulatory specificity depends on where DSX is bound and the ability to coordinate with other sex-, tissue-, or condition-specific factors.

Introduction

In *Drosophila*, sexual dimorphism in the soma is induced by the presence of sex-specific *doublesex* transcripts generated at the end of the sex determination cascade. Female development is initialized by having two X chromosomes, which activates the expression of *Sex-lethal*. Active SXL protein in the female aids in sex-specific splicing of *Transformer* (*tra*) which acts with Transformer-2 (TRA-2) to splice *doublesex* (*dsx)* (Boggs et al., 1987; Nagoshi et al., 1988). TRA-2 is a splicing factor that is expressed in both males and females. However, since only females express active TRA, the TRA/TRA-2 complex mediates splicing of *dsx* transcripts into a female specific mRNA, dsx^F. In the absence of TRA protein in the male, *dsx* transcript gets spliced into the default male specific mRNA, dsx^M (Coschigano et al., 1993; Hertel et al., 1996; Nagoshi et al., 1988). A second factor, *fruitless*, is also sex-specifically spliced by TRA/TRA2 and acts with *dsx* to control sexual dimorphism in the nervous system (Sanders et al., 2008). Thus, the end result of the sex determination cascade is the production of sex specific transcripts for *dsx* and *fru* that encode sex specific products that regulate sexual dimorphism in *Drosophila*.

Based on the little known data on how DSX acts to control sex specific gene expression, the current model for regulation of sexual dimorphism by DSX is that the presence of DSX^F will activate different genes to promote somatic female identity and somatic development, and to also repress male specific genes. On the other hand, DSX^M controls genes that are important for the male developmental program and also to repress female genes required for female identity (Camara et al., 2008; Coschigano et al., 1993). Surprisingly, there are only three known direct targets of DSX (*yp1/2*, *bab1, desatF*), which may not account for the differences in all sexually dimorphic tissues (Burtis et al., 1991; Le Bras et al., 2006; Shirangi et al., 2009; Williams et al., 2008).

In the past decade, few comprehensive approaches have been used to identify genes that are sex-differentially expressed and regulated by the sex-determination hierarchy in adults. Studies primarily using microarrays and serial analysis of gene expression (SAGE) techniques have examined mRNA profiles of tissues (heads and CNS) from flies mutant for *tra*, *dsx* and *fru* at different developmental stages (pupae and adults). Although microarray and SAGE techniques have identified sex-differentially expressed genes of tissues from pupae and adults, these techniques were unable to discover genes deployed through the action of *dsx* in adults (Arbeitman et al., 2004). Moreover, these studies have utilized null mutants of *tra* and *dsx* that have undergone sex reversal and those genes identified as potentially being regulated as a consequence of *tra* or *dsx* could be altered as a consequence of secondary effects.

To get a comprehensive understanding of biologically significant targets of DSX, we combined the dataset produced from our occupancy approaches (Chapter 2) with an RNA-seq dataset from adult flies that have undergone a switch in DSX isoform for an acute period of time. By using conditional mutants of genes involved in sex determination, we can temporally control which DSX isoform will be produced and examine mRNA profiles of gonads from adult females that have undergone the switch from DSX^F to DSX^M (and vice versa) to identify those genes that are being regulated through the action of sex-specific isoforms of DSX. Knowledge of the genes controlled by DSX will elucidate how sexual dimorphism is created and provide insight into the mechanism of DSX function in sexually dimorphic tissues such as the fat body and the gonad.

Materials and Methods

Fly stocks

Fly stocks were obtained from the Bloomington Drosophila Stock Center (Cook et al., 2010). See FlyBase for gene and allele descriptions (Marygold et al., 2013) for *tra2^{ts2}* (FBal0017028), *tra2^{ts1}* (FBal0017027), *P{UAS-tra.F}20J7* (FBti0010566), *P{tubP-GAL80ts}7* (FBti0027798), *P{tubP*-GAL4*}LL7* (FBti0012687), and *w^{1118}*. Flies were grown on standard Bloomington Drosophila Stock Center (Bloomington, IN, USA) medium at 25°C unless otherwise noted.

RNA-seq

Fat body tissue was dissected from age-matched adult flies of the genotypes w^{1118}; *tra2^{ts2}*/ *tra2^{ts1}* (experimental) or *w^{1118}* (control for *dsx^{F}*->dsx^{M} experiments); for *dsx^{M}->dsx^{F}* experiments the genotypes were: *y^{1} w*; P{w^{+mc}=UAS-Tra.F}20J7; P{w^{+mc}=tubP-GAL80^{ts}}7/P{w^{+mc}=tubP-*GAL4*}LL7* (experimental) or *P{w^{+mc}=tubP-GAL80^{ts}}7/P{w^{+mc}=tubP-*GAL4*}LL7* (control). All samples were raised at 18°C until 5 days after eclosion when adults were shifted to either 29°C (for *tra2^{ts}*) or 30°C (for *UAS-TraF*) for 0, 12, or 24 hours. Total RNA was extracted from fat body dissected at room temperature (placed on ice after 30 minutes) using TRIzol Reagent following manufacturer's protocol (Ambion Life Technologies, Carlsbad, CA, USA). Purified RNA was treated with DNAse I following manufacturer's protocol (New England Biolabs, Ipswich, MA, USA) and purified again using phenol:chloroform extraction followed by ethanol precipitation. Duplicate RNA-seq libraries were constructed from 200ng total RNA from independent dissection of each sample using the TruSeq RNA Sample Preparation v2 high-

throughput (HT) protocol (Illumina, San Diego, CA, USA, 2011). Libraries were sequenced on the HiSeq 2000 machine following a 76 bp single-end protocol (Illumina, San Diego, CA, USA).

Reads were generated using the Illumina pipeline software 12.4.2 for all samples excluding control male t=24hr replicate 1 which used pipeline 1.13.48 (re-sequenced due to poor initial sequence quality). Reads passing the Illumina chastity filter were mapped to the *D. melanogaster* genome and assigned to gene models using Tophat 1.4.1 (Trapnell et al., 2009) with a gtf file provided (-G, FlyBase r5.46, see below) and default settings except for the following; minimum intron length was set to 42bp (-i 42) and the maximum multihits was set to 1 (-g 1). Transcript abundance was determined using Cufflinks 2.1.1 (Trapnell et al., 2013) with maximum bundle fragments set to 10,000,000 (--max-bundle-frags 10000000) due to high read density at the *Yp* loci, and upper quartile normalization was used (-N).

To generate a gtf file for Tophat and Cufflinks analyses, the FlyBase GFF annotations (release 5.46) were downloaded from FlyBase as a GFF3 format file. This file was filtered to remove any features on chromosomes Uextra or dmel_mitochondrion_genome as well as the following feature types: enhancer, regulatory_region, exon_junction, rescue_fragment, sequence_variant, pcr_product, point_mutation, orthologous_region, TF_binding_site, protein, chromosome, uncharacterized_change_in_nucleotide_sequence, origin_of_replication, chromosome_band, tandem_repeat, insulator, polyA_site, deletion, BAC_cloned_genomic_insert, complex_substitution, RNAi_reagent, transposable_element_insertion_site, repeat_region, oligonucleotide, breakpoint, transposable_element, chromosome_arm, protein_binding_site, orthologous_to, silencer, region, insertion_site, mature_peptide, DNA_motif, syntenic_region. A leading "chr" was prepended to each chromosome name for consistency with the genomic assembly sequence files used. The

filtered GFF file was imported into a sqlite3 database using gffutils (https://github.com/daler/gffutils), which represents the hierarchical relationships between features as defined in GFF files. For each gene, the "child" transcripts were retrieved from the database, and for each transcript, each child that was either an exon or CDS was retrieved. For each of these exon and CDS features, the gene ID, gene name, transcript ID, and transcript type information were attached to the feature, and it was exported as a GTF format line. The resulting GTF file of exon and CDS features was then run through the gffread program (part of the cufflinks suite) as the command "gffread -E $infile -T -F -o- > $outfile" in order to confirm that the file contained no errors that would prevent downstream use by Cufflinks (Trapnell et al., 2012).

Background expression levels were estimated based on reads in intergenic space (Zhang et al., 2010). Genomic regions that are not located within an annotated gene (FlyBase 5.46), nor within +/- 500 bp flanking an annotated gene, were binned into 199 bp windows (= median of all *D. melanogaster* exons), and FPKM values for these intergenic bins were calculated using the Tophat/Cufflinks parameters used for genes. To prevent loss of mapping between bins, the original intergenic bins were shifted by 100bp and any bin entering a non-intergenic space was removed. The median expression value for all intergenic bins was 1.84839375, and all experimental FPKM values at or below this cutoff were converted to zero. Further, all genes with FPKM=0 in all experimental and control conditions were removed from further analyses.

GEO accession numbers

RNA-seq data are available under GEO series accession GSE49480. Probe sequence, probe position information, and array details are available under GEO accession GPL10639.

GOTerm analyses

Enrichment of gene ontology terms was identified using the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 and AmiGO. The genes from each temperature shift experiment were used as the input dataset. We also used all the genes differentially expressed in each tissue as the input dataset. DAVID returned Enrichment Scores which is a geometric mean (in -log scale) of members p-values in a corresponding annotation culture, used to rank biological significance. Top ranked annotation groups likely have lower p-values for their annotation members. The AmiGO Term Enrichment tool was used to identify what sets of genes have in common by examining annotations and finding significant GO terms shared within a given input dataset. We considered p-values < 0.01 as a significant enrichment.

Statistical analysis

In order to estimate the probability for the same set of genes to appear in two top ranking gene lists between the two studies (occupancy and RNA-seq), we performed a hypergeometric distribution and Fisher's Exact test.

Fisher's Exact test was used in R as follows:

n1=Genes with occupancy, n2=Genes without occupancy, n3=Genes with occupancy in background/control, n4=Genes with occupancy in background/control

FE<-matrix(c(n1, n2, n3, n4), ncol=2, byrow=T)

fisher.test(FE)

Hypergeometric script was used in R as follows:

n=Total genes in Dme genome

n1=Genes with occupancy

n2=Differentially expressed genes

m=Differentially expressed genes with occupancy

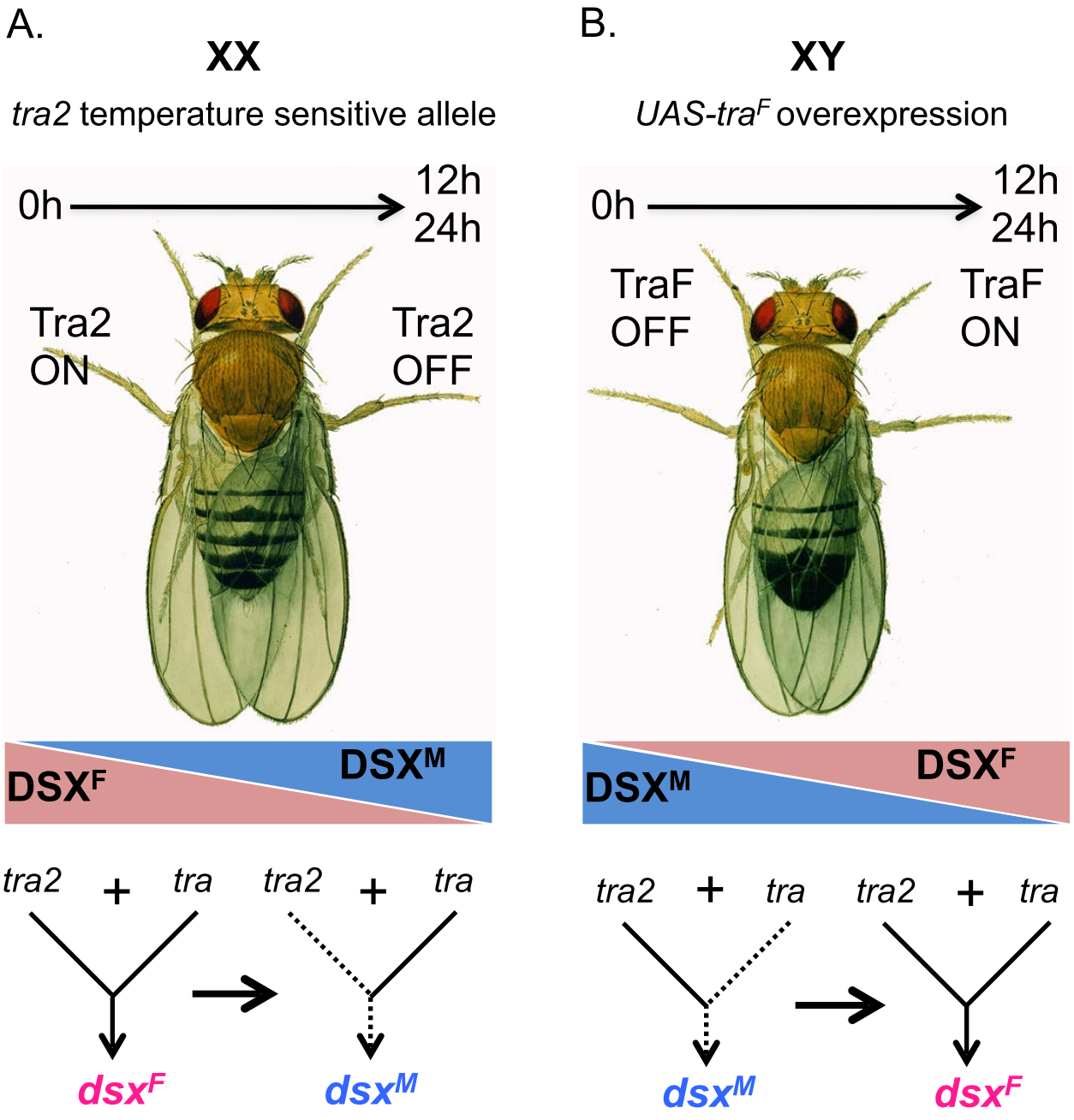
phyper(min(n1, n2), n1, n-n1, n2) - phyper(m-1, n1, n-n1, n2)

Results

RNA-seq on conditional mutants of Transformer

In order to distinguish primary targets of *dsx* and reveal the strategy by which DSX creates sexual dimorphism, we performed RNA-seq on *dsx*-expressing tissues (fat body and gonads) from adult flies that undergone an acute change in DSX isoforms (Figure 3.1). *dsx* expression is under the control of the *Tra* and *Tra2* loci even after sexual differentiation is established (Figure 1.1) (Nagoshi et al., 1988). In XX $tra2^{ts}$ flies maintained at the permissive temperature (18°C), *transformer-2* is functional and *dsx* is spliced into the female specific isoform and flies develop as female. If shifted to the restrictive temperature (29°C), *transformer-2* function is reduced, and *dsx* is spliced into the default male specific form and dsx^{M} is produced. To control for changes in temperature and background, total RNA was isolated from fat body and ovaries from wild-type females (w^{1118}) treated in parallel to XX $tra2^{ts}$ females shifted at the mentioned times. Each time-course experiment on fat body and ovaries was performed in 2 biological replicates.

Figure 3.1. Schematic of RNA-seq paradigm using conditional mutants. A) We performed RNA-seq on dsx-expressing tissues, fat body and ovaries from 4-5 day old adult XX Tra2ts females shifted to 29°C for 0, 12, and 24 hours in order to identify all the genes that are regulated by DSX. To control for changes in temperature and background, total RNA was isolated from fat body and ovaries from wild-type females treated in parallel to XX Tra2ts females shifted at the mentioned times. Each time-course experiment on fat body and ovaries was performed in biological replicate. B) In order to identify the genes regulated by DSXF in the fat body and testis after switching modes of dsx splicing, we ectopically expressed TraF at will in males using the GAL-4/Tubulin-GAL80(ts) system. In order to identify genes regulated through the action of a change in DSX isoform from DSXM to DSXF, we isolated total RNA from male fat body and testes from 4-5 day old adult XY "TraF" males shifted to 30C for 0, 12, and 24 hours. To control for changes in temperature and background, total RNA was also isolated from fat body and testes from GAL-4/Tubulin-GAL80(ts) males treated in parallel to XY "TraF" males shifted at the mentioned times. Each time-course experiment on fat body and testes was performed in biological replicates.

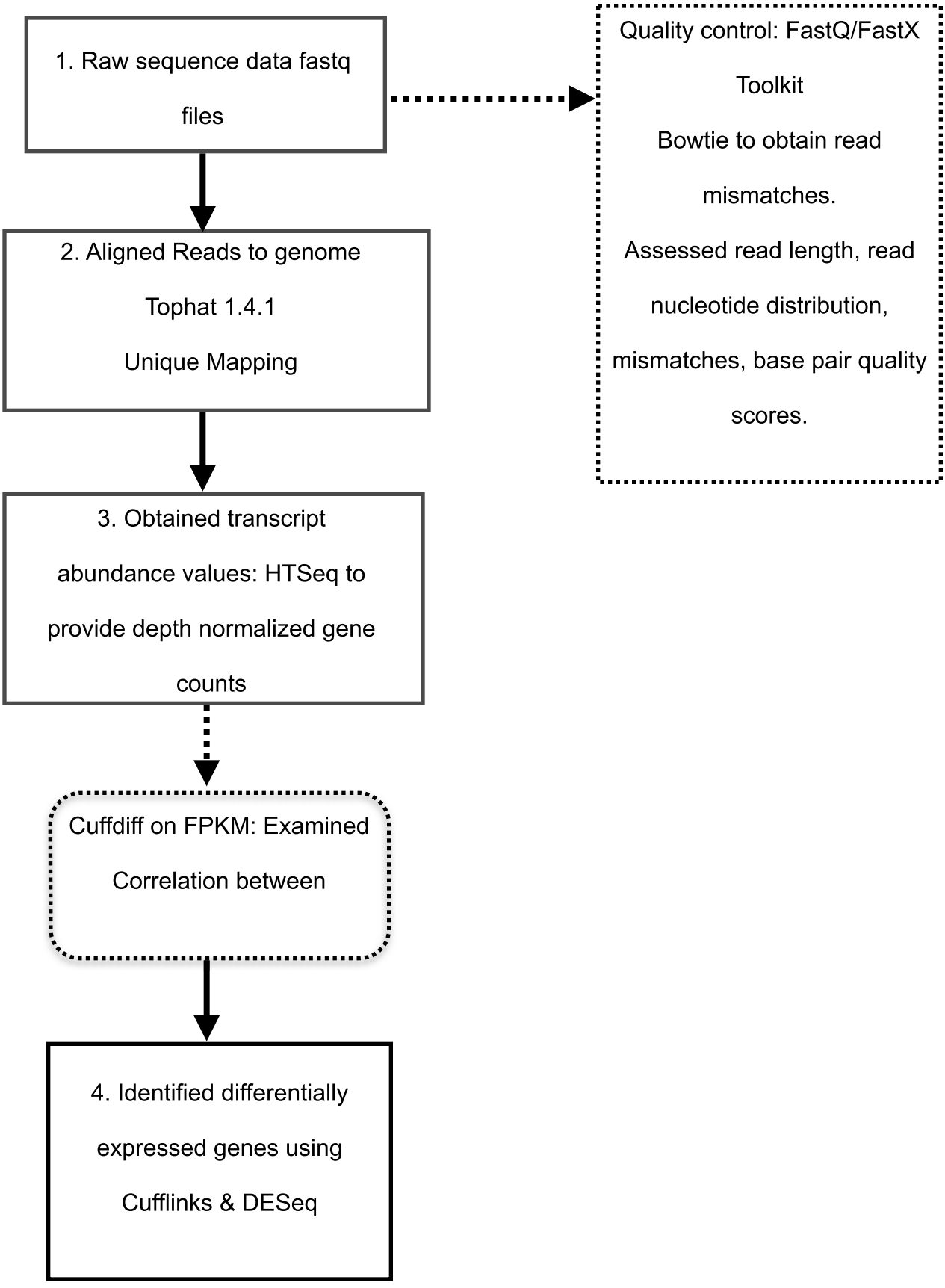

A.
XX
tra2 temperature sensitive allele
0h
12h
24h
Tra2
ON
Tra2
OFF
DSXF
DSXM
tra2 + tra
tra2 + tra
dsxF
dsxM
B.
XY
UAS-traF overexpression
0h
12h
24h
TraF
OFF
TraF
ON
DSXM
DSXF
tra2 + tra
tra2 + tra
dsxM
dsxF

In order to identify genes regulated by DSX in the fat body and testis after switching modes of *dsx* splicing from dsx^M to dsx^F, we ectopically expressed TraF at will in males using the *tubP-GAL80^{ts}*/*tubP*-GAL4 system. We then isolated total RNA from male fat body and testes from 4-5 day old adult XY "TraF" males shifted to 30°C for 0, 12, and 24 hours. To control for changes in temperature and background, total RNA was also isolated from fat body and testes from *tubP-GAL80^{ts}*/*tubP*-GAL4 males treated in parallel to XY "TraF" males shifted at the mentioned times. Each time-course experiment on fat body and testes was performed in 2 biological replicates.

We reasoned that switching between DSX isoform states would provide a greater net change in expression than loss of DSX function since DSX^M and DSX^F are thought to have opposing roles in target gene regulation (Burtis et al., 1991; An and Wensink, 1995). For example, if a gene decreased in expression as a result of a DSX isoform switch from DSX^F to DSX^M, then we would predict that this gene is likely repressed by DSX^M and activated by DSX^F. Alternatively, if a gene increased in expression as a result of a switch in DSX isoform from DSX^M to DSX^F, then we would predict that this gene is likely activated by DSX^F and repressed by DSX^M.

Collectively, we obtained 48 RNA samples, which passed our quality control tests. In collaboration with Dr. Brian Oliver at the NIDDK-NIH, we successfully constructed 48 RNA-seq libraries. Single-end 76 base pair reads were sequenced resulting in a sequencing depth of approximately 30 million reads per sample. For a schematic of the pipeline we used to analyze the RNA-seq data, see Figure 3.2.

Figure 3.2. Overview of the RNA-seq analysis pipeline for detecting differential expression. The steps in the pipeline are in black boxes; the methodological components of the pipeline are in dotted boxes. First raw reads are evaluated for quality and sequence bias; reads are mapped to the Drosophila genome; mapped reads are assembled into expression summaries (via Cufflinks); the data is normalized and statistical testing of differential expression is performed (via Cuffdiff or DESeq).


1. Raw sequence data fastq files
Quality control: FastQ/FastX Toolkit
Bowtie to obtain read mismatches.
Assessed read length, read nucleotide distribution, mismatches, base pair quality scores.
2. Aligned Reads to genome
Tophat 1.4.1
Unique Mapping
3. Obtained transcript abundance values: HTSeq to provide depth normalized gene counts
Cuffdiff on FPKM: Examined Correlation between
4. Identified differentially expressed genes using Cufflinks & DESeq

Mapping and calling differential gene expression

We used FASTQC from the FASTX-Toolkit to assess the quality of raw RNA-seq data and found that the raw sequence data were of high quality and met the FASTQC quality control checks. The FASTX-Toolkit performs quality control checks to spot problems in raw RNA-seq data which originate from either the sequencer or from the starting library material (http://hannonlab.cshl.edu/fastx_toolkit/). First, we analyzed per base sequence quality values across all bases at each position in FastQ files from all 48 samples. Per base sequence quality values were high indicating that base calls were of good quality (Figure 3.3BC). Upon examination of the per base sequence content, we found that in each library there was little to no difference between the different bases of a sequence run.

Random hexamer primers have been shown to cause mismatches in the beginning of the Illumina RNA-seq reads (van Gurp et al., 2013). Thus, we analyzed the mismatch distribution in order to determine if reads required trimming of mismatch bases which would eliminate sequence-specific bias. As expected, the first few basepairs as well as the last basepair did contain high mismatch scores (Figure 3.3A). We trimmed the first 12 and last basepair of all reads followed by mapping of final read alignments having only 2 mismatches. Trimming alone did not increase the percentage of uniquely mapped reads (data not shown).

Additionally, we investigated if we could improve the percentage of uniquely mapped reads by allowing final read alignments to have up to 3 mismatches. We found that the percentage of uniquely mapped reads increased and that the percentage of unmapped reads decreased as the number of mismatches increased. Further, we performed Tophat on reads after trimming poor quality regions and allowing final aligned reads to contain up to 3 mismatches. By combining the two parameters, we found that the percentage of reads uniquely mapped increased

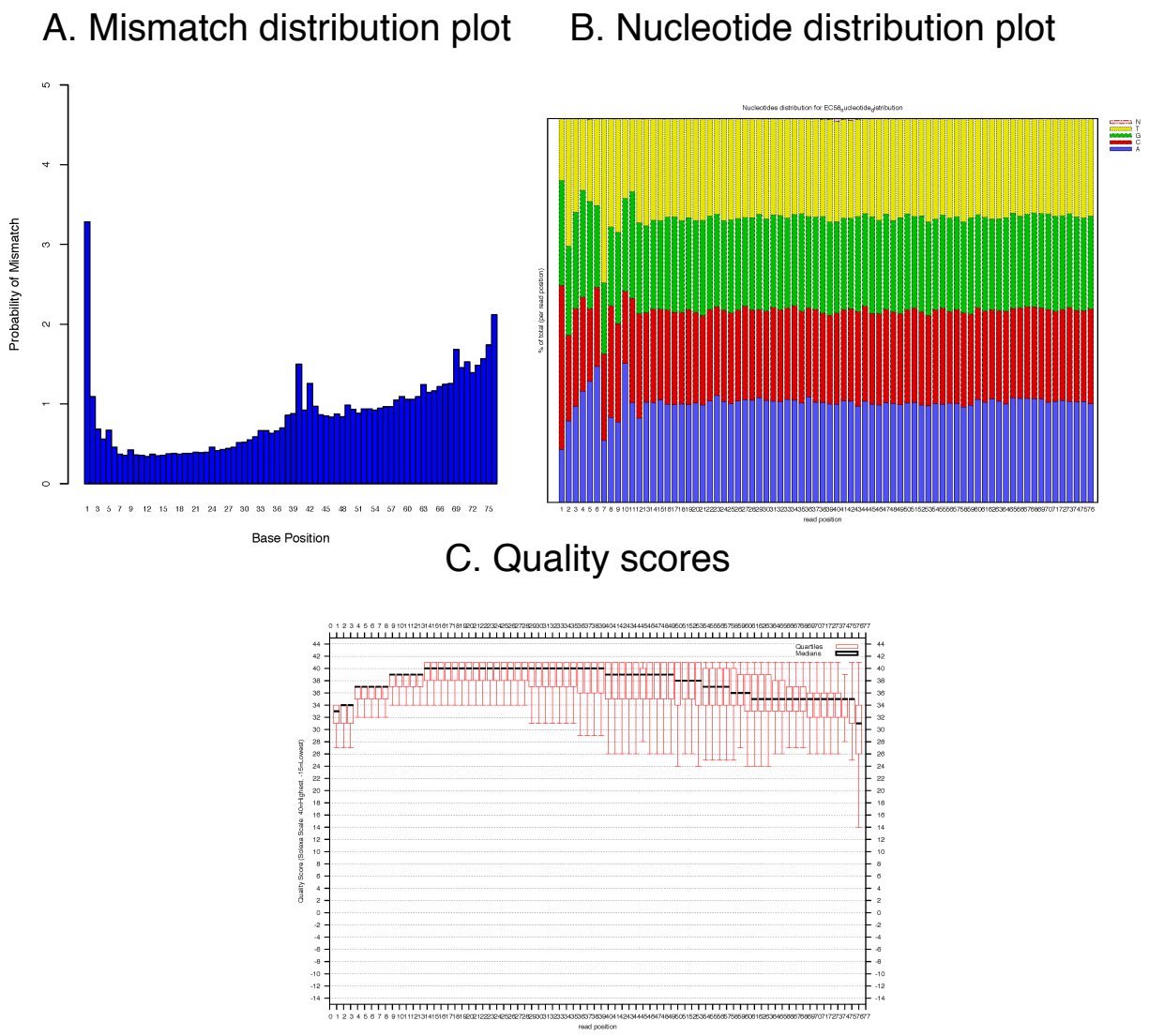


Figure 3.3. Quality control on raw sequence data from RNA-seq via FASTX-Toolkit. The FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) was used to summarize quality scores and nucleotide distributions. For all 48 Samples, mismatch distribution plots show atypical mismatch scores of first few basepairs as well as the last basepair (A). B) The proportion of each base position in a file for which each of the 4 normal DNA bases has been called was plotted out. C) An overview of the range of quality values across all bases at each position in the FastQ file.

by 5%. Although the percentage of uniquely mapped reads improved, the correlation coefficient did not change. Thus, no trimming was required to remove poor quality bases from reads and we continued our analysis on untrimmed reads.

Overall, the raw sequence data met the FastQC quality control checks and were of high quality. Since our reads were of high quality we proceeded to map reads to an annotated Drosophila genome using Tophat. Final read alignments having more than 2 mismatches were discarded. For fat body samples, approximately 73.57% of the total mapped reads were uniquely mapped to the annotated genome, whereas approximately 23.04% fail to map and 0% multimap (Table 3.1). For gonad samples, approximately 58.94% of the total mapped reads were uniquely mapped to the annotated genome, whereas approximately 26.49% fail to map and 0% multimap (Table 3.2). Mapping statistics for each sample are summarized in Table 3.3. Of the few reads that multimapped, they were identified as ribosomal genes or genes that encode tRNAs.

The reproducibility of the results was assessed using a correlation coefficient and shows a high degree of reproducibility among biological replicates. Biological replicates allow for the estimation of within-treatment group variability and provide information that is necessary for making inferences between treatment groups (Auer et al., 2010). Thus, we determined the agreement between biological replicates for gene expression levels and tested the statistical significance of each observed change in expression between them. We found that biological replicates were reproducible and had an average correlation coefficient of 0.97 (Table 3.4, Figure 3.4, Figure 3.5).

In order to identify differentially expressed genes as a consequence of a change in *dsx* isoform, biological replicates were combined and compared to biological treatments using Cufflinks Cuffdiff. First, genes changing in response to a change in temperature were identified

by comparing control treated samples (0 hours compared to 12 hours and 0 hours compared to 24 hours). Genes changing in response to a change in *dsx* isoform and temperature or stress were identified by comparing between treatments (0 hours compared 12 hours and 0 compared 24). By comparing differentially expressed genes in mutant treated samples to controls, genes changing in response to temperature were removed and only genes changing in response to *dsx* remained. A gene was considered differentially expressed if the corrected p-value (q-value) was less than 0.05. In addition to using a stringent corrected p-value cut-off, only genes with a 2-fold or greater change in gene expression were used for subsequent analyses. To ensure that we were observing genes with robust expression changes, we filtered each dataset by requiring an FPKM greater than equal to 1 in at least one sample between compared treatments. After applying all parameters, we found that a significant number of genes were differentially expressed in response to a change in DSX isoform. These results are summarized in Figure's 3.6-3.9 and Table 3.5.

To robustly determine differentially expressed genes, we assessed the impact of using different software packages on identifying differential gene expression in the fat body and gonads. Cufflinks Cuffdiff utilizes a beta negative binomial distribution and DESeq uses a negative binomial distribution. In comparison to DESeq, Cufflinks Cuffdiff is the more conservative pipeline. We found that the genes identified with the more conservative pipeline were typically detected in the less conservative ones. For instance, the lowest number of genes among the fat body datasets was identified with Cufflinks Cuffdiff and at least 80% of these genes were detected with DESeq. In general, DESeq shows an increase in the number of detections compared to Cufflinks Cuffdiff, but both software packages result in overlapping genes differentially expressed (Table 3.7). Since Cufflinks Cuffdiff was the most conservative

software package, we continued our analysis of differential gene expression using the more stringent pipeline.

Table 3.1. RNA-seq mapping statistics for adult fat body. Mapping statistics for all biological replicates in the experiment, including the number of reads, percentage of uniquely mapped reads, percentage of reads that fail to map and multi-mapped reads.

Sample Name	Total # reads	# uniquely mapped	% uniquely mapped	# reads fail to map	% reads fail to map	# multi-mapped	% multi-mapped
FB tra2ts0 R1	27650675	18995965	68.7	8654683	31.3	27	0
FB tra2ts12 R1	21469954	21469954	65.37	11373940	34.63	38	0
FB tra2ts24 R1	24261666	15217596	62.72	9044015	37.28	55	0
FB Wt0 R1	31211821	23024900	73.77	8186856	26.23	65	0
FB Wt12 R1	25289532	17604098	69.61	7685405	30.39	29	0
FB Wt24 R1	32164048	23254073	72.3	8909930	27.7	45	0
FB tra2ts0 R2	19549022	12054280	61.66	7494712	38.34	30	0
FB tra2ts12 R2	30866817	21932819	71.06	8933967	28.94	31	0
FB tra2ts24 R2	23338865	15443621	66.17	7895230	33.83	14	0
FB Wt0 R2	23476506	14103876	60.08	9372587	39.92	43	0
FB Wt12 R2	25807801	25807801	64.03	14495609	35.97	28	0
FB Wt24 R2	26220770	17753560	67.71	8467202	32.29	8	0
FB TraF0 R1	22786414	15436114	67.74	7350255	32.26	45	0
FB TraF12 R1	8308442	4641255	55.86	3667185	44.14	2	0
FB TraF24 R1	28268275	20009194	70.78	8259058	29.22	23	0
FB Control0 R1	28364535	20155544	71.06	8208894	28.94	97	0
FB Control12 R1	28187859	18731369	66.45	9456468	33.55	22	0
FB Control24 R1	10406460	3939109	37.85	6467343	62.15	8	
FB Control24 R1	30720599	13111245	42.68	17609328	57.32%	26	0
FB TraF0 R2	20074865	13988542	69.68	6086286	30.32	37	0
FB TraF12 R2	30814739	21617299	70.15	9197386	29.85	54	0
FB TraF24 R2	28616259	16241753	56.76	12374419	43.24	87	0
FB Control0 R2	26928799	19145325	71.1	7783388	28.9	86	0
FB Control12 R2	26045769	17643669	67.74	8402032	32.26	68	0
FB Control24 R2	33763270	24914607	73.79	8848620	26.21	43	0
FB Control24 R2	99226737	69968759	70.51	29257963	29.49	15	0

Table 3.2. RNA-seq mapping statistics for adult gonads. Mapping statistics for all biological replicates in the experiment, including the number of reads, percentage of uniquely mapped reads, percentage of reads that fail to map and multi-mapped reads.

Sample Name	Total # reads	# uniquely mapped	% uniquely mapped	# reads fail to map	% reads fail to map	# multi-mapped	% multi-mapped
G tra2ts0 R1	36779843	31038022	84.39	5741696	15.61	125	0
G tra2ts12 R1	26781771	22424815	83.73	4356920	16.27	36	0
G tra2ts24 R1	25643883	21615018	84.29	4028835	15.71	30	0
G Wt0 R1	28131498	23526751	83.63	4604737	16.37	10	0
G Wt12 R1	28918534	24154640	83.53	4763874	16.47	20	0
G Wt24 R1	30099130	21346923	70.92	8752116	29.08	91	0
G tra2ts0 R2	28934674	23264269	80.4	5670252	19.6	153	0
G tra2ts12 R2	31837511	26277971	82.54	5559438	17.46	102	0
G tra2ts24 R2	29724271	22035042	74.13	7689199	25.87	30	0
G Wt0 R2	30042353	23625414	78.64	6416917	21.36	22	0
G Wt12 R2	33047111	27310960	82.64	5735927	17.36	224	0
G Wt24 R2	31028019	25592512	82.48	5435408	17.52	99	0
G TraF0 R1	20522701	15750672	76.75	4771988	23.25	41	0
G TraF12 R1	33618876	25569325	76.06	8049485	23.94	66	0
G TraF24 R1	40385391	31060580	76.91	9324799	23.09	12	0
G Control0 R1	29250633	22511980	76.96	6738628	23.04	25	0
G Control12 R1	25103216	19242152	76.65	5861049	23.35	15	0
G Control24 R1	30566932	23767160	77.75	6799677	22.25	95	0
G TraF0 R2	28193536	22549807	79.98	5643708	20.02	21	0
G TraF12 R2	32064543	25774817	80.38	6289695	19.62	31	0
G TraF24 R2	31175574	24983971	80.14	6191561	19.86	42	0
G Control0 R2	35993404	28621264	79.52	7372106	20.48	34	0
G Control12 R2	25380580	20601865	81.17	4778702	18.83	13	0
G Control24 R2	21821438	17406027	79.77	4415392	20.23	19	0

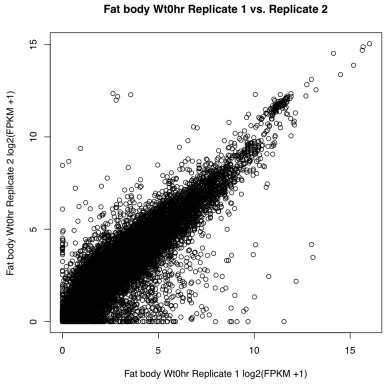
Table 3.3. RNA-seq mapping statistics summary for adult fat body and gonads. Mapping statistics for all biological replicates in the experiment, including averages and standard deviation of the number of reads, percentage of uniquely mapped reads, percentage of reads that fail to map and multi-mapped reads.

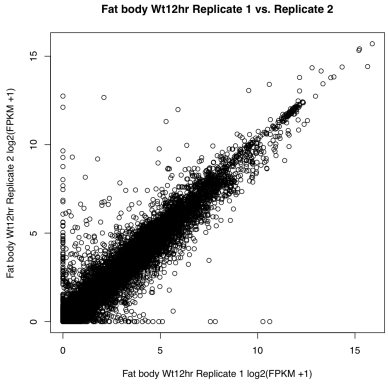
	Total # reads	# uniquely mapped	% uniquely mapped	# reads fail to map	% reads fail to map	# multi-mapped	% multi-mapped
Average	27982070.88	20685105.31	67.72	7462718.58	24.42	50.85	0.00
Standard Deviation	14239642.20	10552360.61	20.02	4469053.94	10.68	43.24	0.00
Fat body Average	30143312.24	23141705.32	73.57	7125176.62	23.04	54.51	0.00
Fat body Standard Deviation	16269098.40	11540290.23	14.94	4942220.40	6.89	52.57	0.00
Gonad Average	24740208.86	17000205.31	58.94	7969031.51	26.49	45.35	0.00
Gonad Standard Deviation	10132288.97	7821109.84	23.72	3744344.43	14.71	23.88	0.00

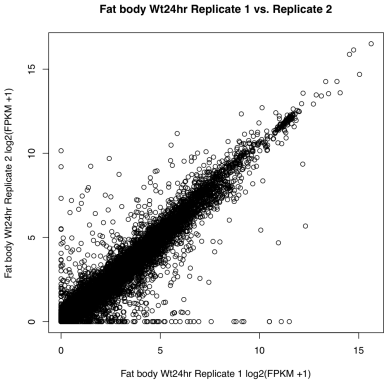
Table 3.4. *log2* transformation of FPKM values and correlation calculation. Transcript abundance values (*f*ragments *p*er *k*ilobase of exon per *m*illion, *fpkm*) were obtained for each sample (via Cufflinks). Pairwise comparisons between biological replicates were performed to test the statistical significance of each observed change in expression between them. The reproducibility of the results was assessed using a correlation coefficient.

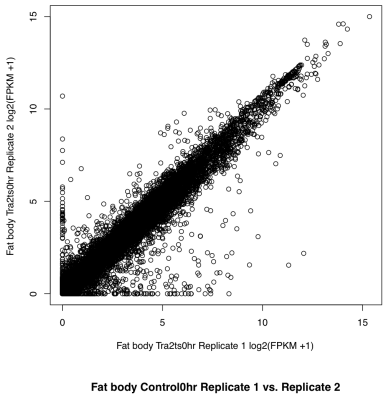
Fat body samples	Correlation coefficient	Gonad samples	Correlation coefficient
tra2ts0 rep1 vs rep2	0.9643283	tra2ts0 rep1 vs rep2	0.9816857
tra2ts12 rep1 vs rep2	0.9321906	tra2ts12 rep1 vs rep2	0.9826624
tra2ts24 rep1 vs rep2	0.9725335	tra2ts24 rep1 vs rep2	0.9658641
wt0 rep1 vs rep2	0.9192168	wt0 rep1 vs rep2	0.9768281
wt12 rep1 vs rep2	0.9586364	wt12 rep1 vs rep2	0.9771535
wt24 rep1 vs rep2	0.9601666	wt24 rep1 vs rep2	0.9776805
traF0 rep1 vs rep2	0.9631531	traF0 rep1 vs rep2	0.9854376
traF12 rep1 vs rep2	0.8910348	traF12 rep1 vs rep2	0.9816894
traF24 rep1 vs rep2	0.9397956	traF24 rep1 vs rep2	0.9880533
control0 rep1 vs rep2	0.9650104	control0 rep1 vs rep2	0.9451565
control12 rep1 vs rep2	0.9405382	control12 rep1 vs rep2	0.981091
control24 rep1 vs rep2	0.9539249	control24 rep1 vs rep2	0.9881425
Average Correlation Coefficient Value	*1*	*Average Correlation Coefficient Value*	*1*
Standard deviation	*0*	*Standard deviation*	*0*

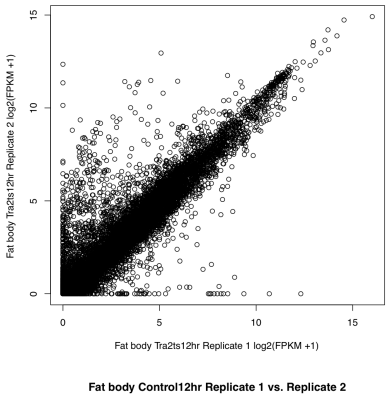
Figure 3.4. Log-base mean-variance correlation between technical and biological replicates for adult fat body. Scatterplots of *log* mean expression values (FPKM) against the *log* of the variance across technical and biological replicates at the transcript and gene levels.

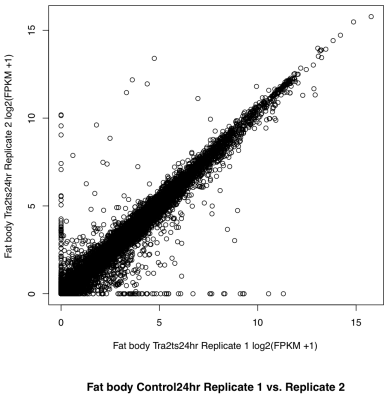

Fat body Wt0hr Replicate 1 vs. Replicate 2
Fat body Wt0hr Replicate 2 log2(FPKM +1)
Fat body Wt0hr Replicate 1 log2(FPKM +1)

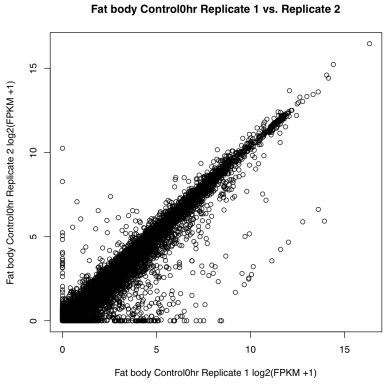

Fat body Wt12hr Replicate 1 vs. Replicate 2
Fat body Wt12hr Replicate 2 log2(FPKM +1)
Fat body Wt12hr Replicate 1 log2(FPKM +1)

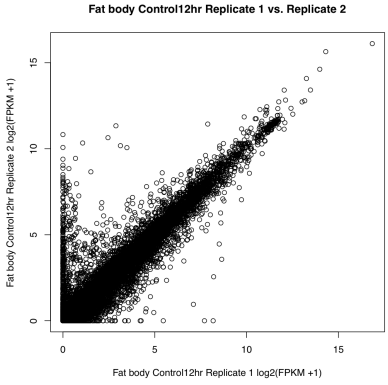

Fat body Wt24hr Replicate 1 vs. Replicate 2
Fat body Wt24hr Replicate 2 log2(FPKM +1)
Fat body Wt24hr Replicate 1 log2(FPKM +1)

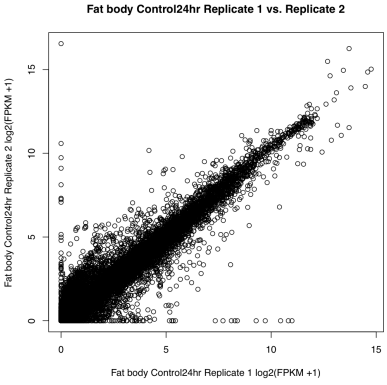

Fat body Tra2ts0hr Replicate 1 vs. Replicate 2
Fat body Tra2ts0hr Replicate 2 log2(FPKM +1)
Fat body Tra2ts0hr Replicate 1 log2(FPKM +1)

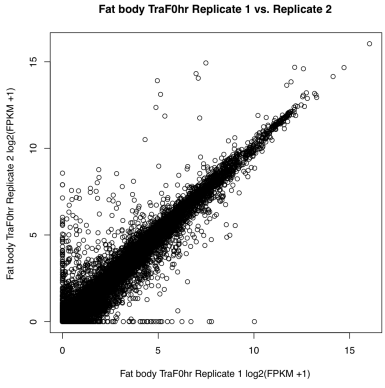

Fat body Tra2ts12hr Replicate 1 vs. Replicate 2
Fat body Tra2ts12hr Replicate 2 log2(FPKM +1)
Fat body Tra2ts12hr Replicate 1 log2(FPKM +1)

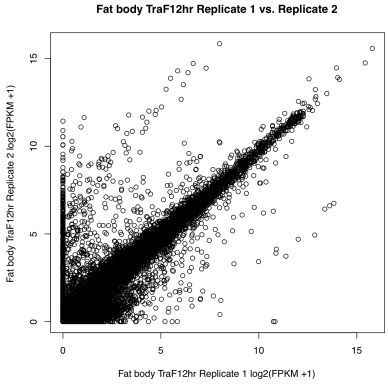

Fat body Tra2ts24hr Replicate 1 vs. Replicate 2
Fat body Tra2ts24hr Replicate 2 log2(FPKM +1)
Fat body Tra2ts24hr Replicate 1 log2(FPKM +1)


Fat body Control0hr Replicate 1 vs. Replicate 2
Fat body Control0hr Replicate 2 log2(FPKM +1)
Fat body Control0hr Replicate 1 log2(FPKM +1)


Fat body Control12hr Replicate 1 vs. Replicate 2
Fat body Control12hr Replicate 2 log2(FPKM +1)
Fat body Control12hr Replicate 1 log2(FPKM +1)


Fat body Control24hr Replicate 1 vs. Replicate 2
Fat body Control24hr Replicate 2 log2(FPKM +1)
Fat body Control24hr Replicate 1 log2(FPKM +1)


Fat body TraF0hr Replicate 1 vs. Replicate 2
Fat body TraF0hr Replicate 2 log2(FPKM +1)
Fat body TraF0hr Replicate 1 log2(FPKM +1)


Fat body TraF12hr Replicate 1 vs. Replicate 2
Fat body TraF12hr Replicate 2 log2(FPKM +1)
Fat body TraF12hr Replicate 1 log2(FPKM +1)


Fat body TraF24hr Replicate 1 vs. Replicate 2
Fat body TraF24hr Replicate 2 log2(FPKM +1)
Fat body TraF24hr Replicate 1 log2(FPKM +1)

Figure 3.5. Log-base mean-variance correlation between technical and biological replicates for adult gonads. Scatterplots of *log* mean expression values (FPKM) against the *log* of the variance across technical and biological replicates at the transcript and gene levels.

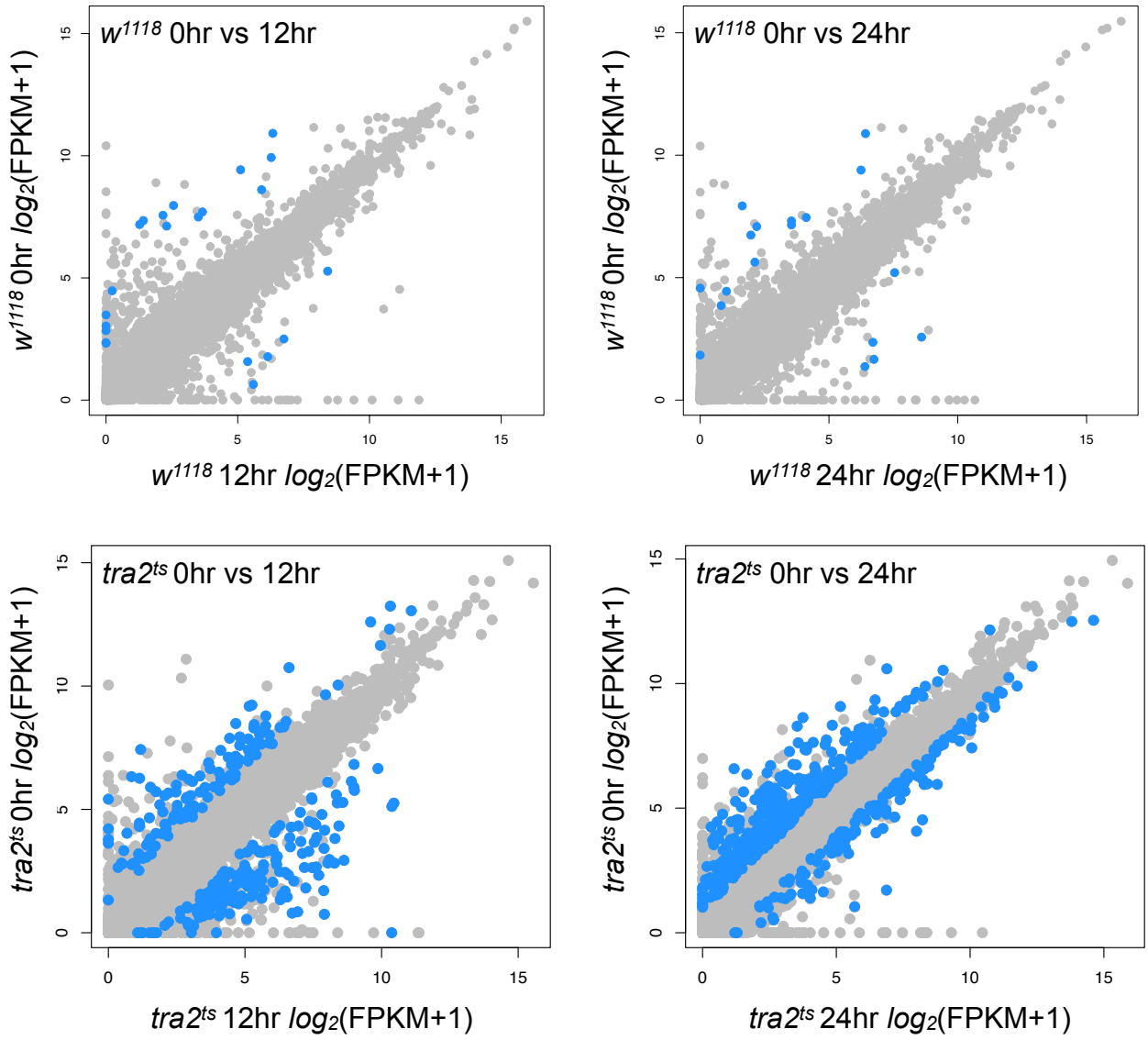
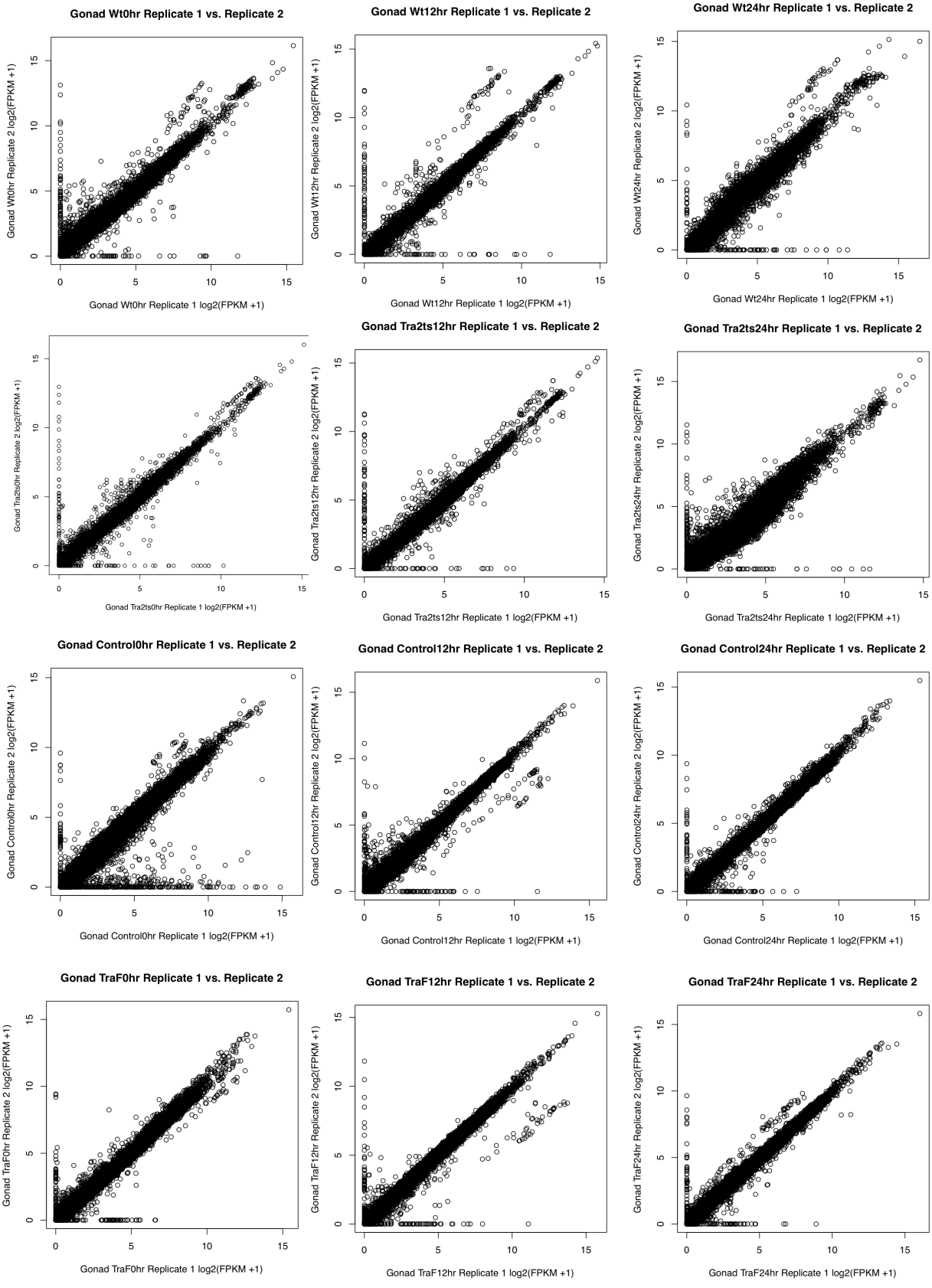


Figure 3.6. Scatter plots of differential gene expression in female fat body of *tra2*^{ts} and *w*^{1118} 12hrs and 24hrs following DSX isoform shift. Scatter plots of FPKM between time points in log scale. Light grey, annotated Drosophila transcripts; blue, differentially expressed genes following indicated temperature shift.


Gonad Wt0hr Replicate 1 vs. Replicate 2
Gonad Wt0hr Replicate 2 log2(FPKM +1)
Gonad Wt0hr Replicate 1 log2(FPKM +1)
Gonad Wt12hr Replicate 1 vs. Replicate 2
Gonad Wt12hr Replicate 2 log2(FPKM +1)
Gonad Wt12hr Replicate 1 log2(FPKM +1)
Gonad Wt24hr Replicate 1 vs. Replicate 2
Gonad Wt24hr Replicate 2 log2(FPKM +1)
Gonad Wt24hr Replicate 1 log2(FPKM +1)
Gonad Tra2ts0hr Replicate 2 log2(FPKM +1)
Gonad Tra2ts0hr Replicate 1 log2(FPKM +1)
Gonad Tra2ts12hr Replicate 1 vs. Replicate 2
Gonad Tra2ts12hr Replicate 2 log2(FPKM +1)
Gonad Tra2ts12hr Replicate 1 log2(FPKM +1)
Gonad Tra2ts24hr Replicate 1 vs. Replicate 2
Gonad Tra2ts24hr Replicate 2 log2(FPKM +1)
Gonad Tra2ts24hr Replicate 1 log2(FPKM +1)
Gonad Control0hr Replicate 1 vs. Replicate 2
Gonad Control0hr Replicate 2 log2(FPKM +1)
Gonad Control0hr Replicate 1 log2(FPKM +1)
Gonad Control12hr Replicate 1 vs. Replicate 2
Gonad Control12hr Replicate 2 log2(FPKM +1)
Gonad Control12hr Replicate 1 log2(FPKM +1)
Gonad Control24hr Replicate 1 vs. Replicate 2
Gonad Control24hr Replicate 2 log2(FPKM +1)
Gonad Control24hr Replicate 1 log2(FPKM +1)
Gonad TraF0hr Replicate 1 vs. Replicate 2
Gonad TraF0hr Replicate 2 log2(FPKM +1)
Gonad TraF0hr Replicate 1 log2(FPKM +1)
Gonad TraF12hr Replicate 1 vs. Replicate 2
Gonad TraF12hr Replicate 2 log2(FPKM +1)
Gonad TraF12hr Replicate 1 log2(FPKM +1)
Gonad TraF24hr Replicate 1 vs. Replicate 2
Gonad TraF24hr Replicate 2 log2(FPKM +1)
Gonad TraF24hr Replicate 1 log2(FPKM +1)

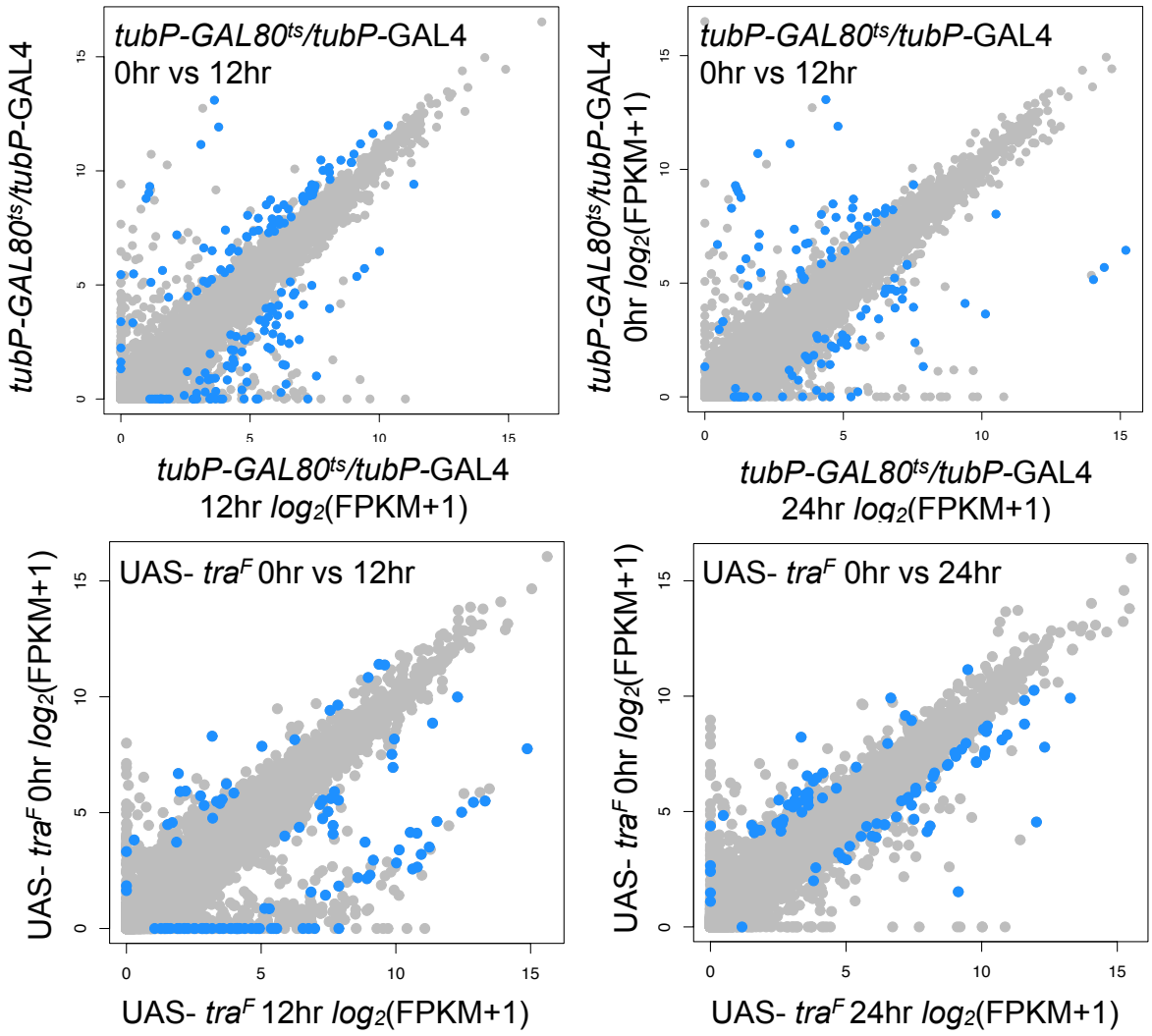


Figure 3.7. Scatter plots of differential gene expression in male fat body of UAS- *tra*^{F} and *tubP-GAL80*^{ts}*/tubP*-GAL4 control 12hrs and 24hrs following DSX isoform shift. Scatter plots of FPKM between time points in log scale. Light grey, annotated Drosophila transcripts; blue, differentially expressed genes following indicated temperature shift.

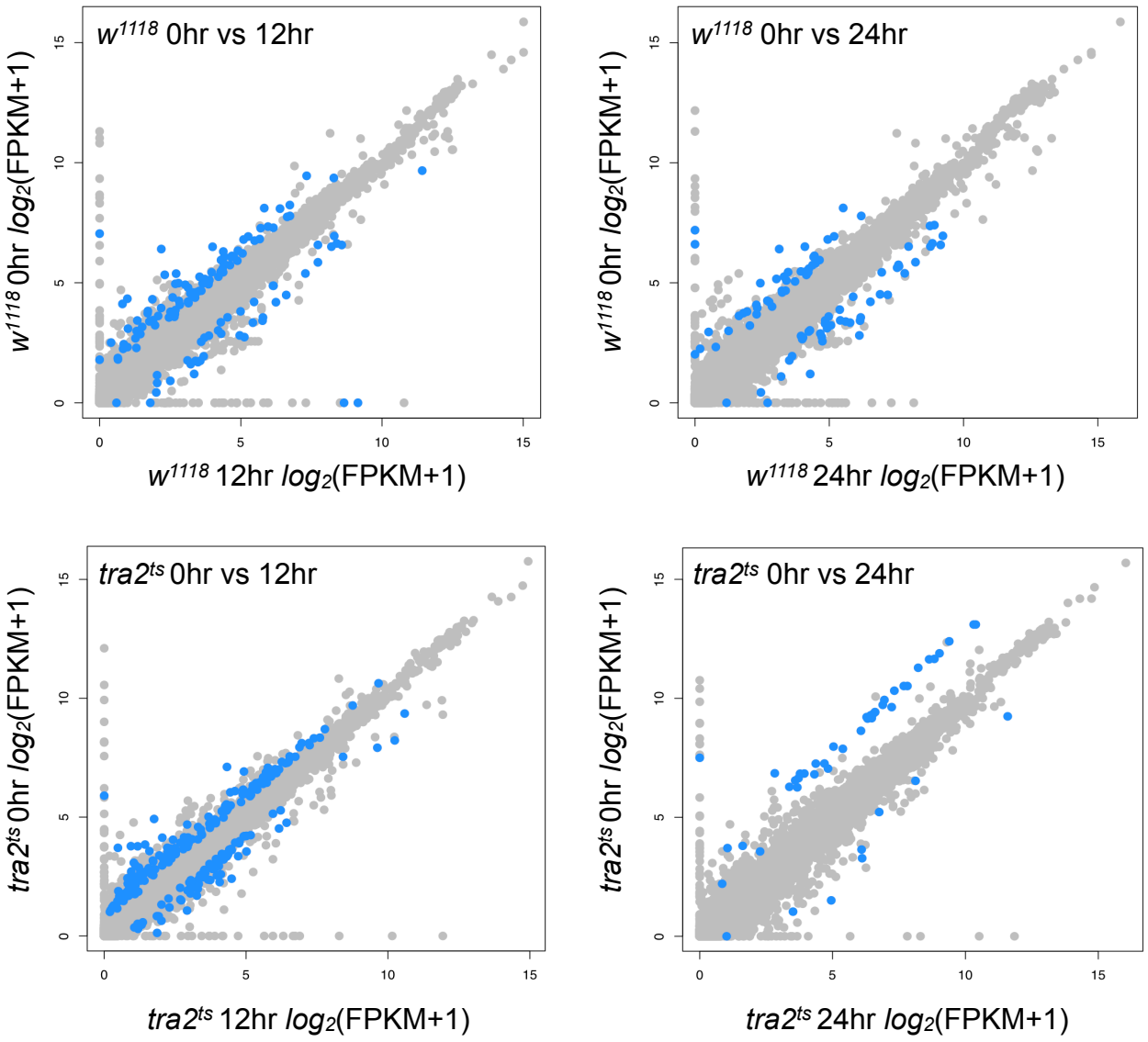


Figure 3.8. Scatter plots of differential gene expression in female gonads of *tra2*^{ts} and *w*^{1118} 12hrs and 24hrs following DSX isoform shift. Scatter plots of FPKM between time points in log scale. Light grey, annotated Drosophila transcripts; blue, differentially expressed genes following indicated temperature shift.

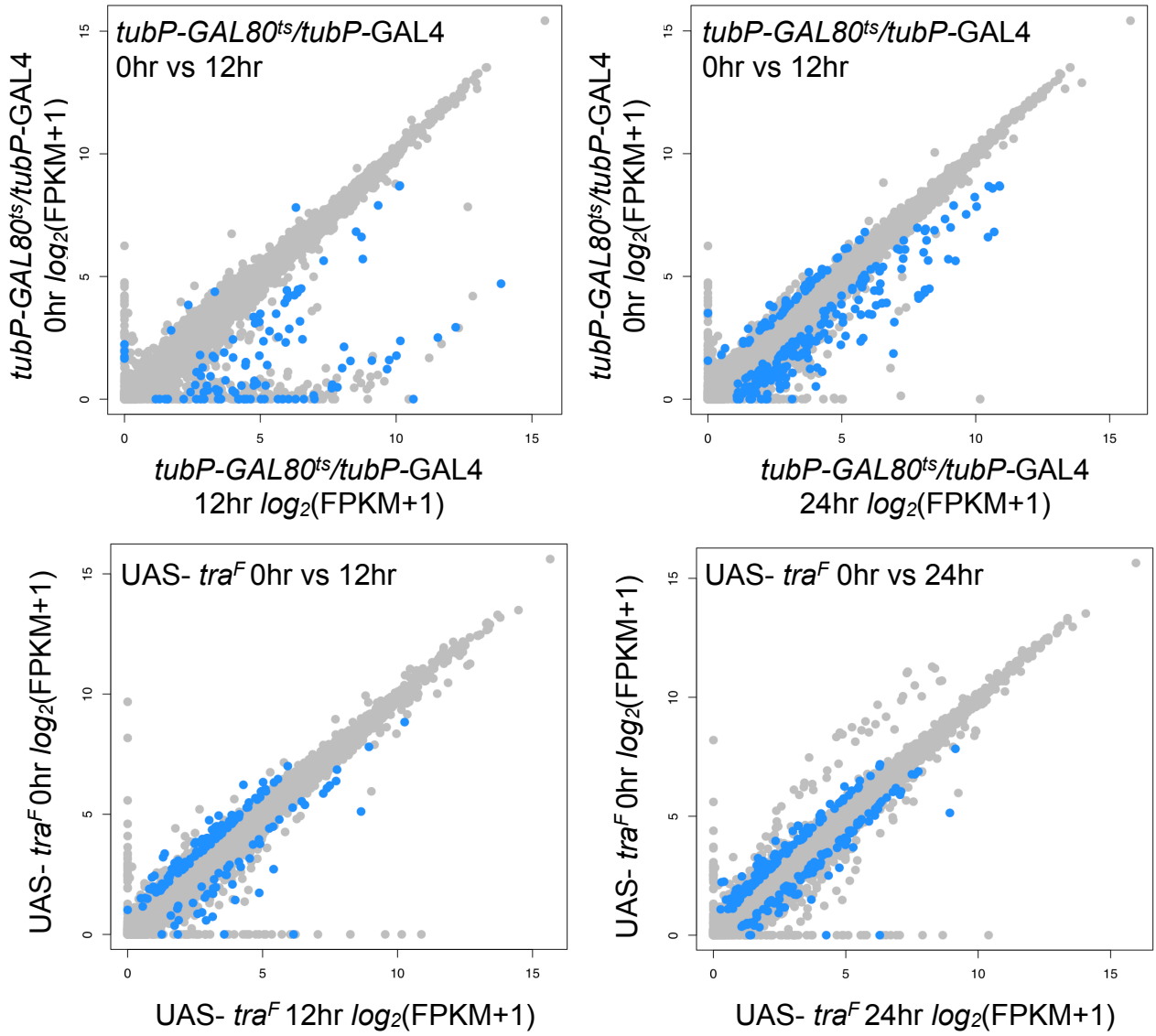


Figure 3.9. Scatter plots of differential gene expression in male gonads of UAS- *tra*^{F} and *tubP-GAL80*^{ts}*/tubP*-GAL4 control 12hrs and 24hrs following DSX isoform shift. Scatter plots of FPKM between time points in log scale. Light grey, annotated Drosophila transcripts; blue, differentially expressed genes following indicated temperature shift.

Table 3.5. Differentially expressed genes in fat body and gonads according to Cufflinks Cuffdiff. Differentially expressed genes were identified in fat body of animals that undergone a switch in DSX isoform and controls. A gene was considered differentially expressed as possessing statistically significant differential expression (adjusted p value <0.05) and required that one FPKM value for either the 0-hour or examined timepjoint be greater than zero.

Sample name	Differentially expressed genes
Female fat body tra2ts12hr	298
Female fat body tra2ts24hr	474
Female fat body wt12hr	22
Female fat body wt24hr	18
Male fat body traF12hr	104
Male fat body traF24hr	91
Male fat body control12hr	167
Male fat body control24hr	113
Female ovary tra2ts12hr	171
Female ovary tra2ts24hr	47
Female ovary wt12hr	133
Female ovary wt24hr	86
Male testis traF12hr	99
Male testis traF24hr	109
Male testis control12hr	107
Male testis control24hr	191

Table 3.6. Differentially expressed genes in fat body and gonads according to DESeq. Differentially expressed genes were identified in fat body of animals that undergone a switch in DSX isoform and controls. A gene was considered differentially expressed as possessing statistically significant differential expression (adjusted p value <0.05) and required that one FPKM value for either the 0-hour or examined timepjoint be greater than zero.

Sample name	Differentially expressed genes
Female fat body tra2ts12hr	610
Female fat body tra2ts24hr	865
Female fat body wt12hr	175
Female fat body wt24hr	139
Male fat body traF12hr	420
Male fat body traF24hr	452
Male fat body control12hr	504
Male fat body control24hr	247
Female ovary tra2ts12hr	143
Female ovary tra2ts24hr	70
Female ovary wt12hr	303
Female ovary wt24hr	264
Male testis traF12hr	115
Male testis traF24hr	140
Male testis control12hr	364
Male testis control24hr	332

Table 3.7. Analysis of differentially expressed genes in fat body and gonads identified using different software packages. Cufflinks Cuffdiff utilizes a beta negative binomial distribution and DESeq uses a negative binomial distribution. In general, DESeq shows an increase in the number of detections compared to Cufflinks Cuffdiff, but both software packages result in overlapping genes differentially expressed.

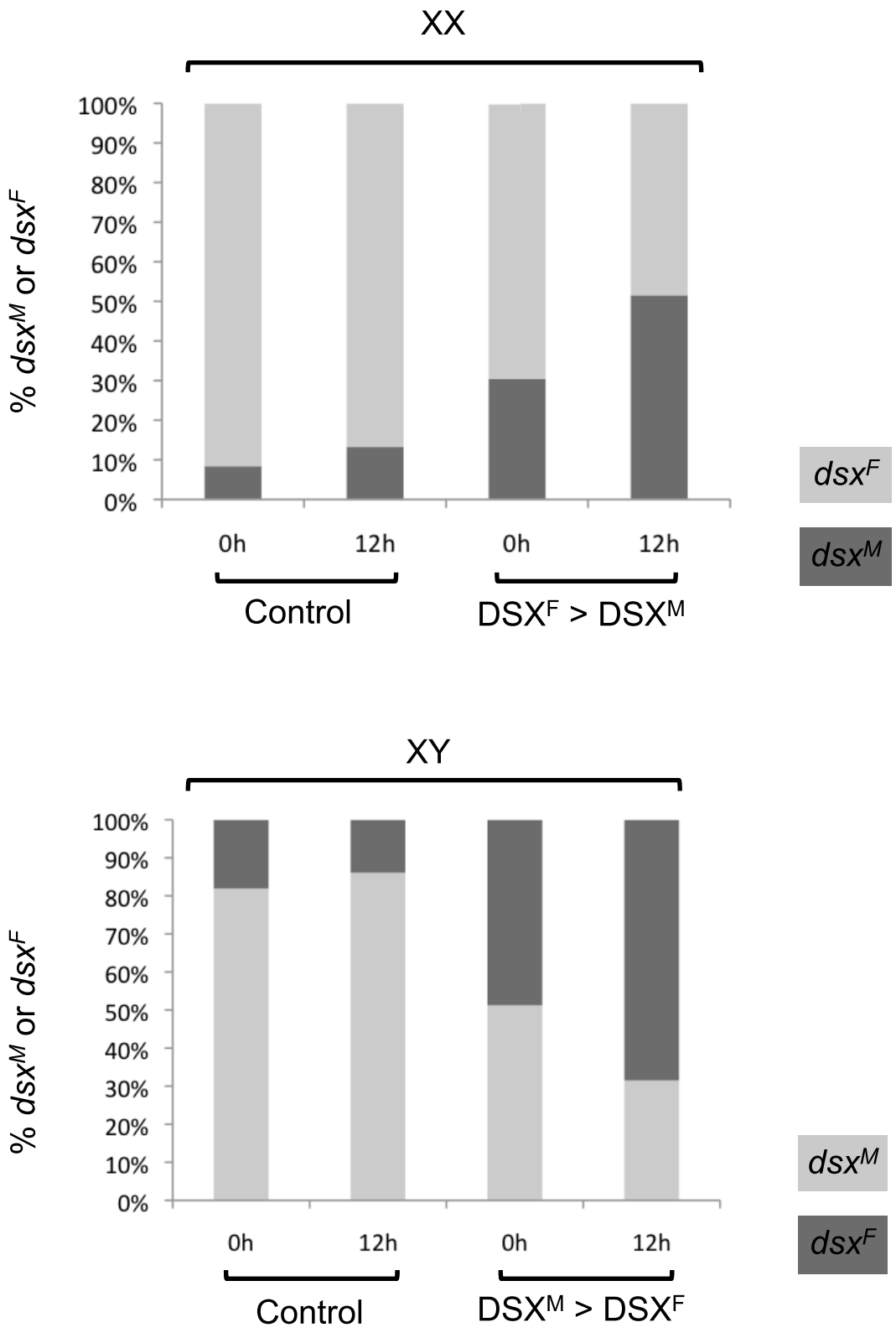
Sample Name	Cuffdiff	DESeq	Cuffdiff genes found in DESeq output (%)
	Differentially expressed genes (adjusted p value<0.05)		
Female fat body tra2ts12hr	298	610	246 (83%)
Female fat body tra2ts24hr	474	865	437 (92%)
Female fat body wt12hr	22	175	22 (100%)
Female fat body wt24hr	18	139	18 (100%)
Male fat body traF12hr	104	420	78 (75%)
Male fat body traF24hr	91	452	84 (92%)
Male fat body control12hr	167	504	154 (92%)
Male fat body control24hr	113	247	92 (81%)
Female ovary tra2ts12hr	171	143	98 (69%)
Female ovary tra2ts24hr	47	70	22 (47%)
Female ovary wt12hr	133	303	113 (85%)
Female ovary wt24hr	86	264	77 (90%)
Male testis traF12hr	99	115	78 (79%)
Male testis traF24hr	109	305	86 (79%)
Male testis control12hr	107	364	95 (89%)
Male testis control24hr	191	332	157 (82%)

dsx isoform-bias and changes in gene expression of known DSX targets

XX *tra2^{ts}* flies were morphologically female and fertile while maintained at 18°C, indicating that sufficient DSX^{F} activity existed to support female-specific development and physiology. Similarly, XY; *tubP-GAL80^{ts}/tubP*-GAL4 flies are phenotypically male and fertile when grown at 18°C. However, when we analyzed splice junction counts of *dsx* using Spanki 0.4.2 (Sturgill et al., 2013) in fat body, we observed some expression of both male and female *dsx* RNA splice forms at the permissive temperature of 18°C in both genotypes. Importantly, upon shifting *tra2^{ts}* females to 29°C for 12 or 24hrs, *dsx* mRNAs encoding the *dsx^{M}* isoform were elevated ~2-fold ($p << 0.01$; Fisher's Exact Test) indicating that we succeeded in manipulating *dsx* isoform-bias in these flies. For the reciprocal acute shift in isoform in UAS-*tra^{F}* males, we found that the portion of *dsx* spliced into the female isoform increased ~ 5 fold ($p << 0.01$; Fisher's Exact Test) (Figure 3.10).

There is no publication reporting direct targets of DSX in the gonad which might be used as positive controls. However, the model described earlier on how DSX isoforms have opposing roles in vivo is based on the observation that in females the DSX^{F} isoform directly activates expression of the YP genes in the fat body (Burtis et al., 1991; An and Wensink, 1995). In males, the DSX^{M} isoform directly represses activation of the YP genes. Thus, the repression and activation of the yolk protein genes in the adult fat body as a consequence of a change in DSX isoform over time is a reporter we used to validate our fat body RNA-seq dataset. Among the genes that change in response to a switch in DSX isoform, from DSX^{F} to DSX^{M}, in the female fat body *tra2^{ts}* RNA-seq dataset were the yolk protein genes (*Yp1/2*) (Figure 3.11). In female fat body undergoing a change in isoform from DSX^{F} to DSX^{M}, both *yolk protein 1* and *2* decrease

Figure 3.10. *dsx* isoform bias in DSX switched animals. Upon shifting $tra2^{ts}$ females to 29°C for 12 or 24hrs, *dsx* mRNAs encoding the dsx^{M} isoform were elevated ~2-fold ($p << 0.01$; Fisher's Exact Test) in adult fat body. For the reciprocal acute shift in isoform in UAS-tra^{F} males, the portion of dsx spliced into the female isoform increased ~ 5 fold ($p << 0.01$; Fisher's Exact Test).


XX
% dsx^M or dsx^F
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
0h
12h
0h
12h
Control
DSX^F > DSX^M
dsx^F
dsx^M
XY
% dsx^M or dsx^F
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
0h
12h
0h
12h
Control
DSX^M > DSX^F
dsx^M
dsx^F

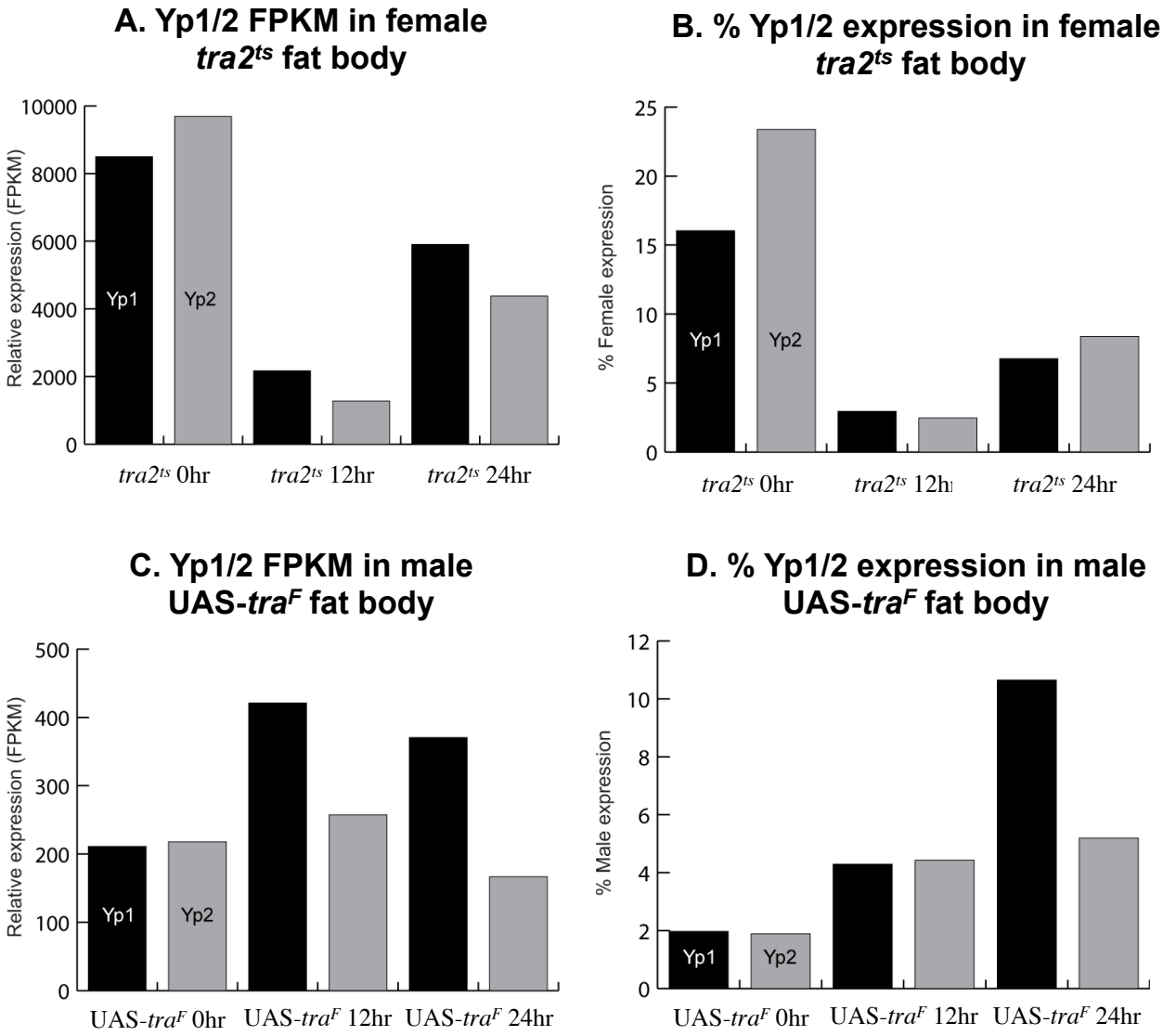


Figure 3.11. Yp1/2 show an increase in expression correlating with higher DSX^F relative to DSX^M and vice versa. A, C). Relative expression of Yp1/2 (FPKM) in $tra2^{ts}$ female and UAS-tra^F male fat body. B, D) Percent Yp1/2 expression in $tra2^{ts}$ female and UAS-tra^F male fat body. The Yp1/2 response was expected based on known DSX regulation, thus confirming that we manipulated known DSX outputs.

by at least 2 fold within 12 hours (p value<0.05). In the 24 hour shifted animals, though not statistically significant, the yolk protein genes are decreasing in expression 2 fold as a consequence of DSX^M expression. In male UAS-tra^F animals shifted from DSX^M to DSX^F, the change in expression of the yolk protein genes were not significant, however the yolk protein gene expression is activated over time as a consequence of DSX^F. The *Yp1/2* response was expected based on known DSX regulation, thus confirming that we manipulated known DSX outputs. In addition to the *Yp1/2* response, we also observed an increase in *Fad2* expression correlating with higher DSX^F relative to DSX^M. The *Fad2* locus encodes a female-specific sterol desaturase involved in sex pheromone signaling (Chertemps et al., 2006) that is directly regulated by DSX in oenocytes (Shirangi et al., 2009). Though not statistically significant, our data indicate that DSX also regulates *Fad2* in the fat body.

Mode of dsx regulation in the fat body and the gonads

Using the genes identified as being differentially expressed after the change in DSX isoforms we identified the direction of gene expression changes using a 2-fold change cut off. A gene was classified as being activated in response to DSX if expression change increased greater by at least 2-fold at 12 hrs and 24 hrs of the shift. A gene was classified as being repressed in response to a DSX if expression change was less by at least 2-fold at 12 hrs and 24 hrs of the shift. Using this criteria, we determined whether target genes favored the current model of DSX action, if DSX^M represses or DSX^F promotes its target expression.

In the fat body, we first examined if differentially expressed genes were shared between treatments. We reasoned that genes responding directly to a change in DSX isoform would provide a greater net change in expression over time and be shared between treatments. In fat

body of *tra2ts* females that undergone a switch in DSX isoform, from DSX^F to DSX^M, we found that 40% of the genes that changed were shared between treatments, 12 hrs and 24 hrs, following temperature shift (Figure 3.13). Upon shifting $tra2^{ts}$ females to 29°C for 12 hrs, 58% of differentially expressed genes were activated and 42% of differentially expressed genes were repressed. By 24 hrs of the shift, we observed that 30% of differentially expressed genes were activated and 70% of differentially expressed genes were repressed. Thus, following the DSX isoform switch, from DSX^F to DSX^M, the amount of genes repressed in response to an increase in DSX^M is elevated (Table 3.8, Figure 3.12).

For the reciprocal acute shift in isoform in UAS-$traF$ male fat body, we found that 22% of the genes that changed were shared between treatments, 12 hrs and 24 hrs, following temperature shift (Figure 3.13). Upon shifting UAS-*traF* males to 30°C for 12 hrs, 75% of differentially expressed genes were activated and 25% of differentially expressed genes were repressed. By 24 hrs of the shift, 55% of differentially expressed genes were activated and 45% of genes were repressed (Table 3.8, Figure 3.12). Though we did not observe a consistent increase in the number of differentially expressed genes activated in response to a switch from DSX^M to DSX^F in both treatments, we found that genes become activated in response to an increase in DSX^F isoform relative to the initial time point, 0 hrs.

In ovaries of *tra2ts* females that have undergone a switch in DSX isoform, from DSX^F to DSX^M, we found that 19% of the genes that changed were shared between treatments, 12 hrs and 24 hrs, following temperature shift (Figure 3.13). Upon shifting $tra2^{ts}$ females to 29°C for 12 hrs, 26% of differentially expressed genes were activated and 74% of differentially expressed genes were repressed. By 24hrs of the shift, we observed that 17% of differentially expressed genes

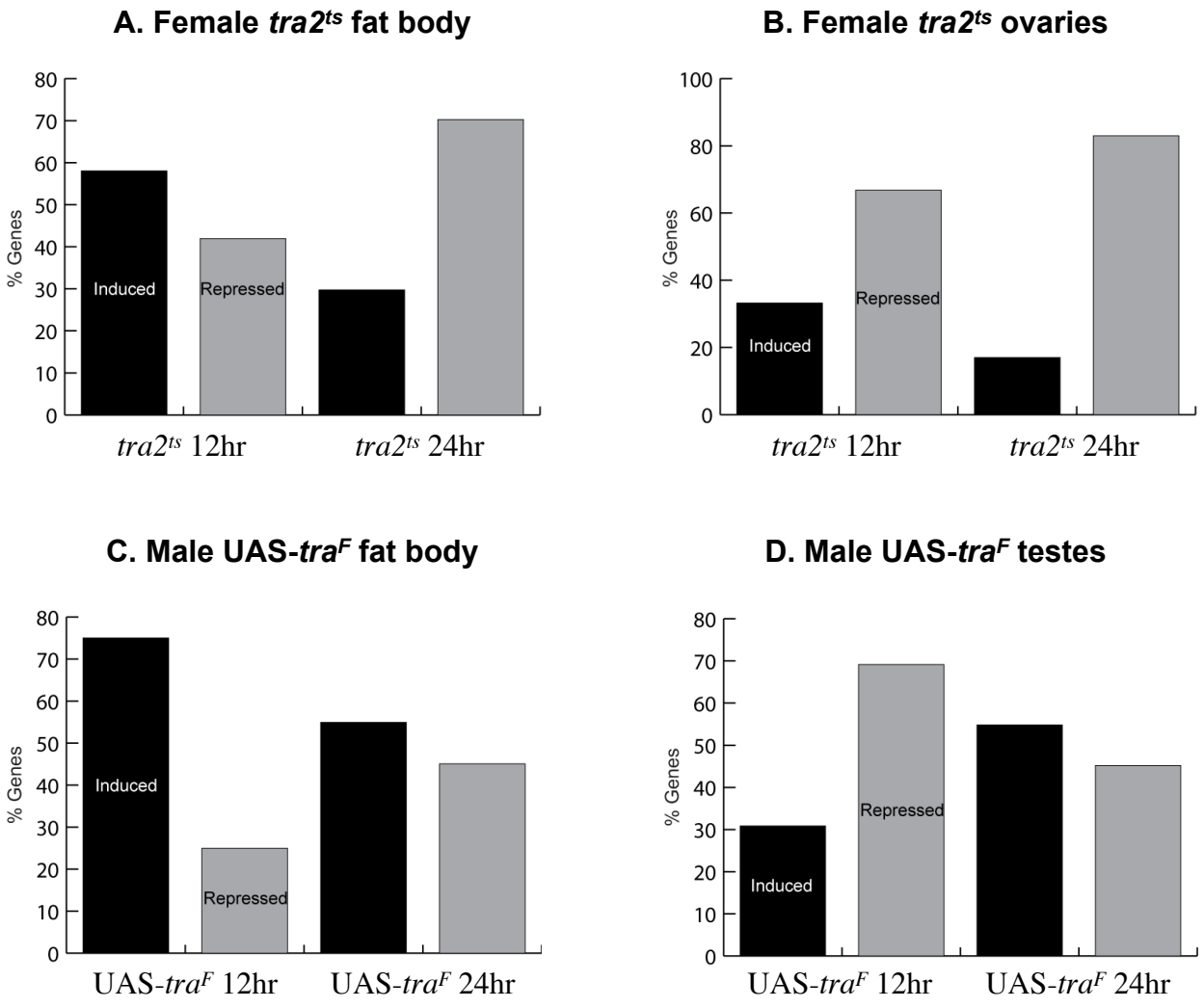


Figure 3.12. The mode of DSX regulation in fat body and gonads. A gene was classified as being activated in response to DSX if expression change increased greater than or equal to 2-fold at 12hrs and 24hrs of the shift. A gene was classified as being repressed in response to a DSX if expression change was less than or equal to 2-fold at 12hrs and 24hrs of the shift. A) Following the DSX isoform switch, from DSXF to DSXM, the amount of genes repressed in response to an increase in DSXM is elevated in fat body (A) and ovaries (B) from tra2^{ts} females. In the opposite experiment going from DSXM to DSXF, the amount of genes activated in response to an increase in DSXF is elevated in fat body (C) and testes (D) from UAS-tra^{F} males.

Table 3.8. Summary of the mode of DSX regulation in fat body and gonads. A gene was classified as being activated in response to DSX if expression change increased greater than or equal to 2-fold at 12hrs and 24hrs of the shift. A gene was classified as being repressed in response to a DSX if expression change was less than or equal to 2-fold at 12hrs and 24hrs of the shift.

Sample Name	Activated	Repressed	% Activated	% Repressed
Female Fat body Tra2ts 12	173	125	58	42
Female Fat body Tra2ts 24	141	333	30	70
Male Fat body TraF 12	78	26	75	25
Male Fat body TraF 24	50	41	55	45
Ovary Tra2ts 12	45	125	26	74
Ovary Tra2ts 24	8	39	17	83
Testis TraF 12	38	60	39	61
Testis TraF 24	56	52	52	48

were activated and 83% of differentially expressed genes were repressed. Following the DSX isoform switch, from DSX^F to DSX^M, the amount of genes repressed in response to an increase in DSX^M is elevated supporting our prediction that DSX regulated genes would have a greater net change in expression over time (Table 3.8, Figure 3.12).

For the reciprocal acute shift in isoform in UAS-tra^F male testes, we found that 31% of the genes that changed were shared between treatments, 12 hrs and 24 hrs, following temperature shift (Figure 3.13). Upon shifting UAS-*traF* males to 30°C for 12hrs, 39% of differentially expressed genes were activated and 61% of differentially expressed genes were repressed. By 24 hrs of the shift, 52% of differentially expressed genes were activated and 48% of genes were repressed. Similar to UAS-tra^F male fat body experiments, we did not observe an increase in the number of differentially expressed genes activated in response to a switch from DSX^M to DSX^F in both treatments. However, our data suggests that genes become activated in response to an increase in DSX^F isoform (Table 3.8, Figure 3.12).

While we do find a modest overlap between treatments after inducing a switch in DSX isoform from DSX^F to DSX^M, and vice versa, each time point has unique transcripts differentially expressed (Figure 3.13). Upon a switch in DSX isoform, a modest number of differentially expressed genes were shared among "activated" and "repressed" genes, indicating that we consistently manipulated gene expression over time and that differential genes are likely DSX dependent in both tissues we examined (Table 3.8, Figure 3.12, 3.13). Further, we find that this switch in DSX isoform indeed provides a greater net change in expression over time.

Since DSX plays opposing roles *in vivo* on known targets, we predicted that genes activated upon switching DSX isoforms from DSX^F to DSX^M and vice versa that we would capture the same genes repressed by DSX^M in $tra2^{ts}$ females and activated by UAS-tra^F. We

compared genes “activated” in *tra2*^{ts} female fat body to genes “repressed” in UAS-*traF* male fat body and found that 33% of these genes overlapped. When we compared genes “repressed” in *tra2*^{ts} females to genes “activated” in UAS-*traF* males we found that only 6% overlapped (Figure 3.14). This data suggests that some genes regulated as a consequence of *dsx* are regulated in the *Yp*-like mode of regulation. In comparison to the fat body, differentially expressed genes responding to a switch in DSX isoform in the gonad appeared to be regulated by one isoform of DSX. When we examined if the same genes repressed by DSX^{M} in *tra2*^{ts} female ovaries were activated by UAS-*tra*^{F} in male testes, we found that only 8% of these genes overlapped. In the opposite comparison, we found that 0% of genes overlapped between genes “activated” in *tra2*^{ts} female ovaries and genes “repressed” in UAS-*tra*^{F} male testes (Figure 3.15). Thus, genes that are regulated as a consequence of *dsx* are not all regulated in the *Yp*-like mode of regulation. Instead, our data suggests that there genes that are regulated downstream of one isoform of DSX.

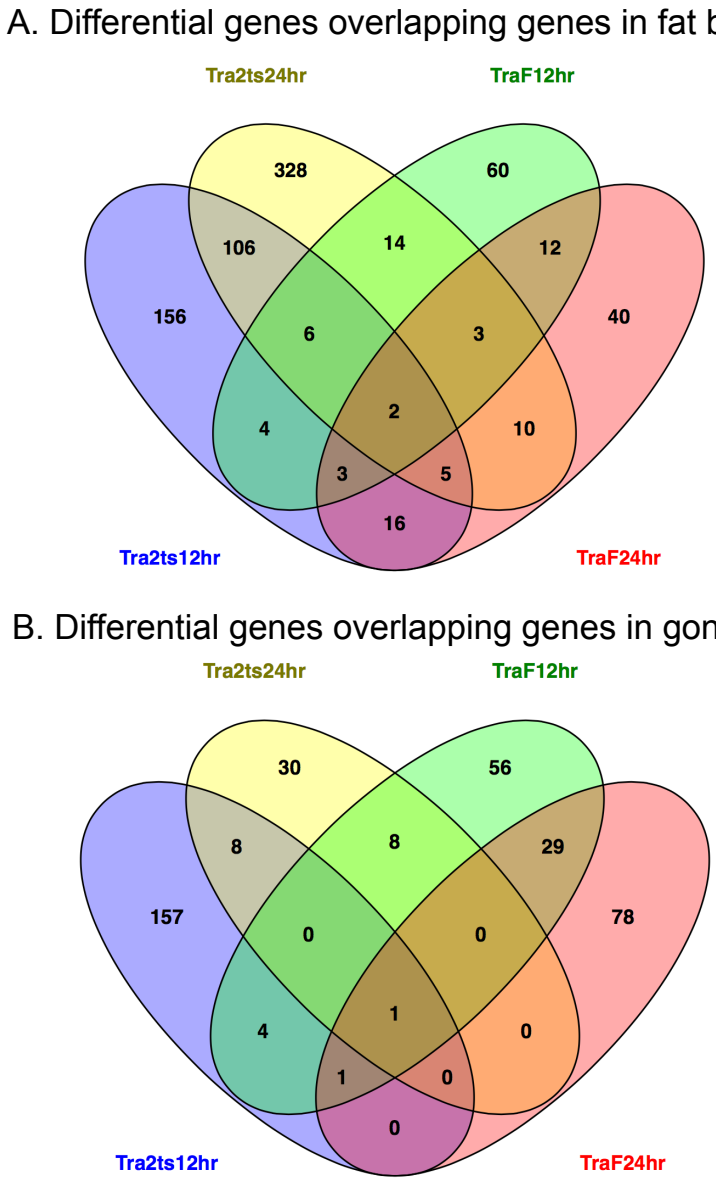


Figure 3.13. Overlap between treatments after inducing a switch in DSX isoform. Differentially expressed genes are shared between 4-, 3-, and 2- experiments as well as each experiment having unique differential transcripts. A) Differentially expressed genes in adult fat body undergoing a switch in DSX isoform from DSX^F to DSX^M, and vice versa. B) Differentially expressed genes in adult gonads undergoing a switch in DSX isoform from DSX^F to DSX^M, and vice versa.

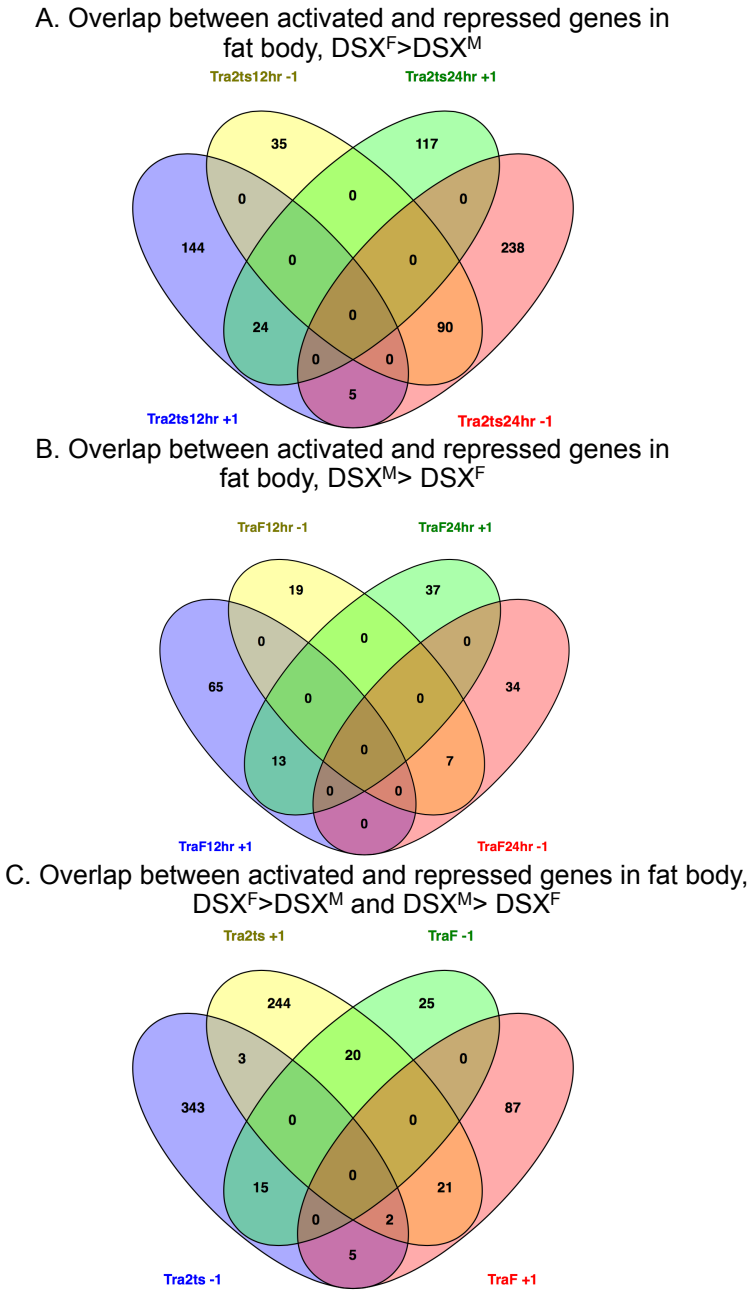


Figure 3.14. Overlap between activated and repressed genes in fat body undergoing a switch in DSX isoform. A) Following a shift in *tra2*^{ts} fat body from DSX^{F}>DSX^{M} and vice versa in UAS-*tra*^{F} fat body DSX^{F}>DSX^{M} (B), a number of genes change in the same direction between the 12hr and 24hr time points for each given genotype. The majority of transcripts are unique to a time point for each experiment. C) genes are not inversely regulated in fat body of *tra2*^{ts} compared to fat body of UAS-*tra*^{F}.

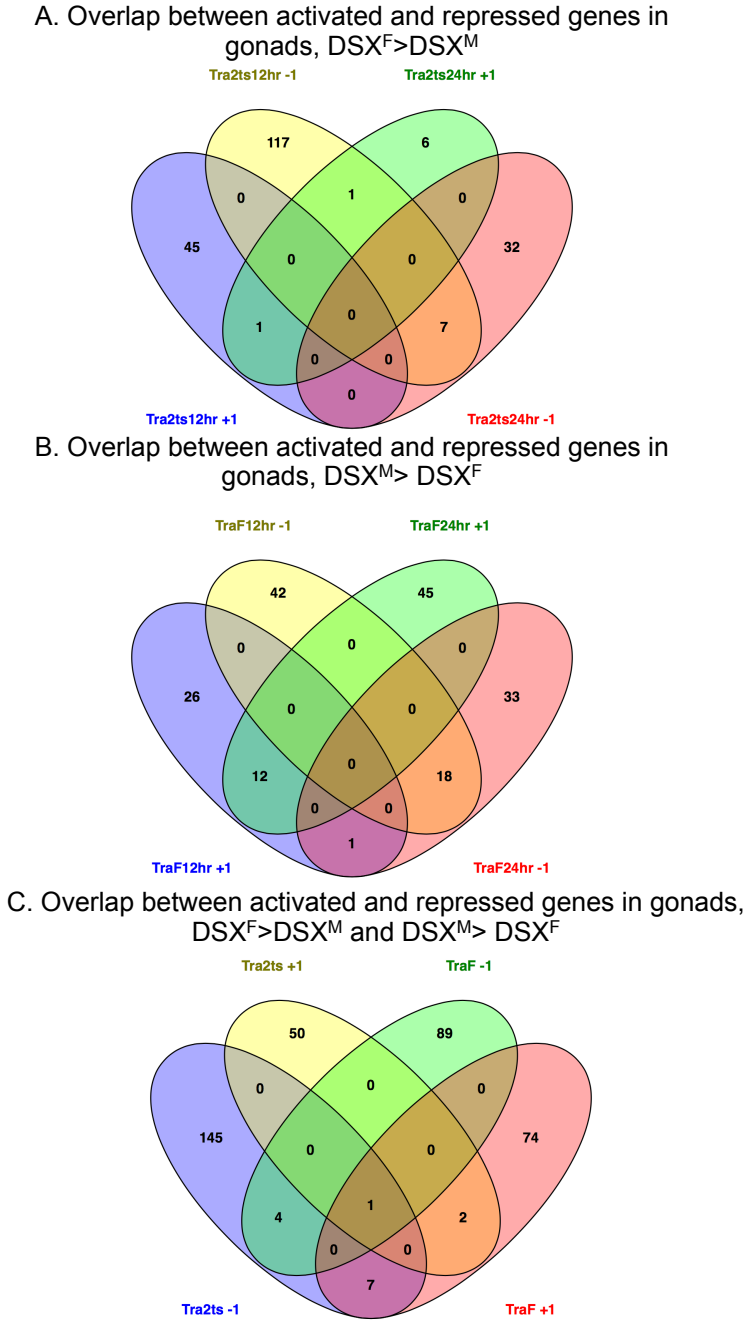


Figure 3.15. Overlap between activated and repressed genes in gonads undergoing a switch in DSX isoform. A) Following a shift in *tra2*^{ts} gonads from DSX^{F}>DSX^{M} and vice versa in UAS-*tra*^{F} DSX^{F}>DSX^{M} (B), few genes change in the same direction between the 12hr and 24hr time points for each given genotype. The majority of transcripts are unique for each experiment. C) genes are not inversely regulated in gonads of *tra2*^{ts} compared to UAS-*tra*^{F} gonads.

Functional analysis of differentially expressed genes

To provide functional interpretation of our whole-transcriptome data, we examined enrichment of gene ontology (GO) terms among differentially expressed genes. We performed DAVID (the Database for Annotation, Visualization and Integrated Discovery, http://david.abcc.ncifcrf.gov) Functional Annotation Clustering in order to report top ranking groups of similar annotations. For each dataset of differentially expressed genes, we identified clusters of statistically over-represented GO terms using DAVID (Table 3.10-3.14). In the fat body, statistically over-represented GO terms identified by DAVID included categories of genes involved in Regulation of Hormone Levels and Hormone Binding, Sex Differentiation, Mating, Gonad and Genitalia Development, as well as Learning and Memory. In the gonad, categories included Aging, Mating and Behavior, Metabolic Processes, Sex Differentiation, and Female Gamete Generation (Table 3.9). In addition, we utilized AmiGO Term Enrichment tool (http://geneontology.org/page/go-enrichment-analysis) to generate a graphical view of hierarchical terms relating to cellular components, biological and molecular processes enriched in each dataset (Figure 3.16-3.19). We found strong enrichment for many different coherent groups of genes in ontologies using both GO term enrichment tools, supporting the idea that DSX controls a wide range of pathways and functions.

Table 3.9. Functional annotation cluster enrichment of Statistically over-represented GO terms, according to DAVID in expression datasets. Gene lists of differential genes from each time point, 12hrs and 24hrs, in each experiment, *tra2*^{ts} and UAS-*tra*^{F}, were combined for either fat body or gonads and analyzed by DAVID. Functional enriched clustered are listed in descending order with highest number being more statistically significant.

	Fat body	
Annotation cluster	Representative annotation terms	Enrichment score
1	EGF-like	3.04
2	Metal-binding/Cytochrome P450	2.59
3	Regulation of hormone levels	2.10
4	Sex differentiation	1.98
5	Mating	1.63
6	Reproductive behavior	1.55
7	Male sex differentiation	1.49
8	Male sterility, NAD-binding	1.39
9	Female gamete generation	1.34
10	Genitalia development	1.30
11	Gonad development	0.86
12	Pigmentation during development	0.74
13	Hormone binding	0.73
14	Circadian rhythm	0.49
15	Learning and memory	0.43
	Gonad	
Annotation cluster	Representative annotation terms	Enrichment score
1	Metal ion binding/Cytochrome P450	3.08
2	Metabolic process	1.72
3	Female gamete generation	1.61
4	Mating behavior	0.65
5	Molting cycle	0.60
6	Aging	0.56
7	Sex differentiation	0.49
8	Metamorphosis	0.32
9	Cell fate determination	0.26
10	Instar larval or pupal development	0.19

Table 3.10. Functional annotation cluster enrichment of Statistically over-represented GO terms, according to DAVID in female fat body expression datasets. Gene lists of differential genes from each time point, 12hrs and 24hrs, in *tra2*^{ts} fat body was analyzed by DAVID.

tra2ts12	Annotation cluster	Representative annotation terms	Enrichment score
	1	Oxygen transporter activity	4.93
	2	CHK kinase-like	4.92
	3	Proteolysis	4.83
	4	Sectretion/signal peptide	3.38
	5	Symporter activity	2.69
	6	Amino acid transport and metabolism	2.14
	7	Response to heat	1.57
	8	Extracellular matrix part	1.28
	9	Vesicle mediated transport	1.23
	10	Integral to membrane	1.20
	11	Startch and sucrose metabolism	1.19
	12	Regulation of hormone levels	1.04

tra2ts24	Annotation cluster	Representative annotation terms	Enrichment score
	1	Insect cuticle protein	6.18
	2	Secreted/signal peptide	5.74
	3	Chitin metabolic process	4.12
	4	EGF-like	3.59
	5	Oxygen transporter activity	3.31
	6	Proteolysis	3.28
	7	Fatty acid metabolic process	3.16
	8	Extracellular matrix	2.80
	9	Starch and sucrose metabolism	1.65
	10	Cytochrome P450/monooxygenase	1.59
	11	Hormone metabolic process	1.52
	12	Muscle myosin complex	1.47
	13	Male sterility, NAD-binding	1.37
	14	Amino acid transport	1.31
	15	Adult behavior	1.24
	16	Stress-induced protein	1.22
	17	Vesicle-mediated transport/endocytosis	1.16
	18	Integral to membrane	1.13
	19	Hormone binding	1.11
	20	Salivary gland development	1.11
	21	Sex differentiation	1.05

Table 3.11. Functional annotation cluster enrichment of Statistically over-represented GO terms, according to DAVID in male fat body expression datasets. Gene lists of differential genes from each time point, 12hrs and 24hrs, in UAS-*tra*^{F} fat body was analyzed by DAVID. Functional enriched clustered are listed in descending order with highest number being more statistically significant.

traF12	Annotation cluster	Representative annotation terms	Enrichment score
	1	Secreted/signal peptide	6.69
	2	immune response	3.70
	3	Peptidase activity	2.46
	4	Reproductive behavior	2.01
	5	Cytochrome P450/Monooxygenase	1.47
	6	metal-binding	1.14
traF24	Annotation cluster	Representative annotation terms	Enrichment score
	1	Symporter activity	1.83
	2	Cytochrome P450/Monooxygenase	1.45
	3	Peptidase activity/proteolysis	1.32
	4	Secreted/signal peptide	1.13

Table 3.12. Functional annotation cluster enrichment of Statistically over-represented GO terms, according to DAVID in 12hr ovary expression datasets. Gene lists of differential genes from each time point, 12hrs and 24hrs, in *tra2*^{ts} gonad was analyzed by DAVID. Functional enriched clustered are listed in descending order with highest number being more statistically significant.

tra2ts12hr	Annotation cluster	Representative annotation terms	Enrichment score
	1	Secreted signal	2.89
	2	Metal ion binding	2.84
	3	Extracellular matrix	2.29
	4	Actin cytoskeleton	1.83
	5	Histone modification	1.61
	6	Zinc finger	1.40
	7	Glutathione S-transferase	1.30
	8	Endopeptidase activity	1.26
	9	Vitamin binding	1.26
	10	Female gamete generation	1.24
	11	Immune response	0.92
	12	Protein dimerization	0.90
	13	Iron ion binding	0.74
	14	Response to radiation	0.53
	15	Chitin binding	0.47
	16	DNA binding	0.39
	17	Cell fate determination	0.37
	18	Imaginal disc pattern formation	0.26
	19	Ion transport	0.25

Table 3.13. Functional annotation cluster enrichment of Statistically over-represented GO terms, according to DAVID in 12hr testis expression datasets. Gene lists of differential genes from 12hrs in UAS-*tra*^{F} gonad was analyzed by DAVID. Functional enriched clustered are listed in descending order with highest number being more statistically significant.

traF12hr	Annotation clus	Representative annotation terms	Enrichment sco
	1	Amino acid transmembrane transporter activity	4.09
	2	Carboxylesterase activity	2.70
	3	Unknown function	1.77
	4	Transmembrane	1.38
	5	Metal ion binding	1.29
	6	Glutathione S-transferase	1.15
	7	Chitin binding	1.13
	8	Immune response	0.98
	9	Symporter activity	0.71
	10	Structural constituent of cuticle	0.59
	11	Synapse	0.56
	12	Response to radiation	0.51
	13	Glycosylation	0.47
	14	Macromolecular subunit	0.41
	15	Receptor	0.34
	16	RNA splicing	0.28
	17	Metal ion transport	0.26
	18	Plasma membrane	0.26
	19	Ion channel	0.25
	20	Cell-cell signaling	0.22
	21	Cell morphogenesis	0.04
	22	Nuclear lumen	0.03
	23	Nucleotide binding	0.01
	24	DNA-binding	0.01

Table 3.14. Functional annotation cluster enrichment of Statistically over-represented GO terms, according to DAVID in 24hr testis expression dataset. Gene lists of differential genes from 24hrs in UAS-*tra*^{F} gonad was analyzed by DAVID. Functional enriched clustered are listed in descending order with highest number being more statistically significant.

traF24hr	Annotation cluste	Representative annotation terms	Enrichment score
	1	Metal ion binding	1.80
	2	Fatty acid metabolism	1.08
	3	Cell-cell signaling	1.06
	4	Mating behavior	1.05
	5	Membrane	1.04
	6	Cell projection	0.96
	7	Aging	0.88
	8	Metal ion binding	0.86
	9	Immunoglobulin-like	0.85
	10	MAPK signaling pathway	0.76
	11	GTPase regulator activity	0.76
	12	Gene silencing by RNA	0.76
	13	Structural constituent of cuticle	0.68
	14	Cell morphogenesis	0.66
	15	Immune response	0.61
	16	Enzyme inhibitor activity	0.58
	17	mRNA processing	0.44
	18	Secreted signal	0.41
	19	Cell motion	0.39
	20	Cell death	0.36
	21	Axis specification	0.33
	22	Imaginal disc pattern formation	0.33
	23	Receptor	0.31
	24	Organelle membrane	0.26
	25	Instar larval or pupal development	0.23
	26	Nucleotide binding	0.04

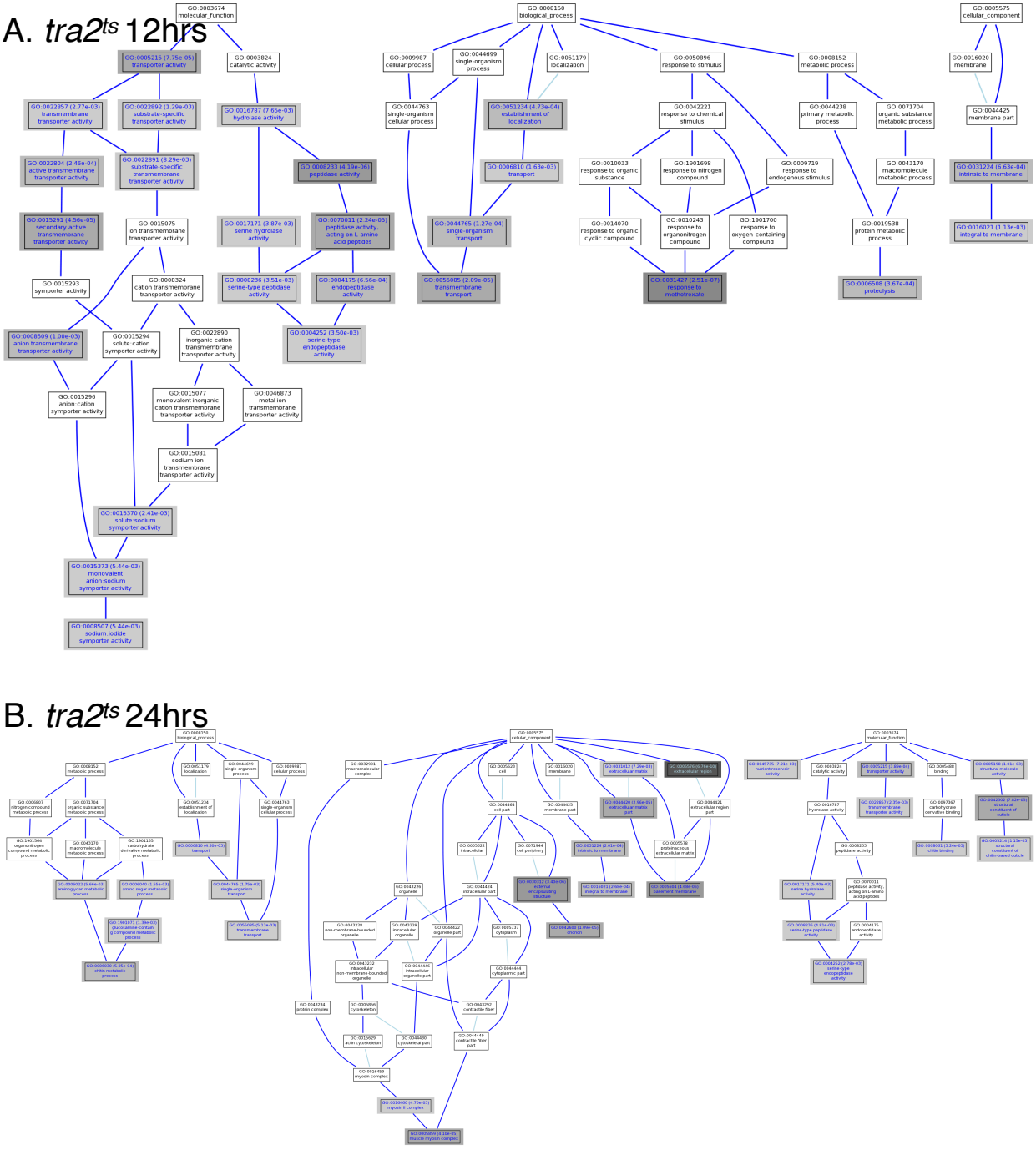


Figure 3.16. Graphical view of hierarchical terms relating to cellular components, biological and molecular processes enriched in *tra2*^{ts} fat body 12hrs (A) and 24hrs (B) following temperature shift. Terms are related to each other in a semi-hierarchical fashion (a directed graph structure), from very broad terms (at the top of the hierarchy) to specific (at the bottom).

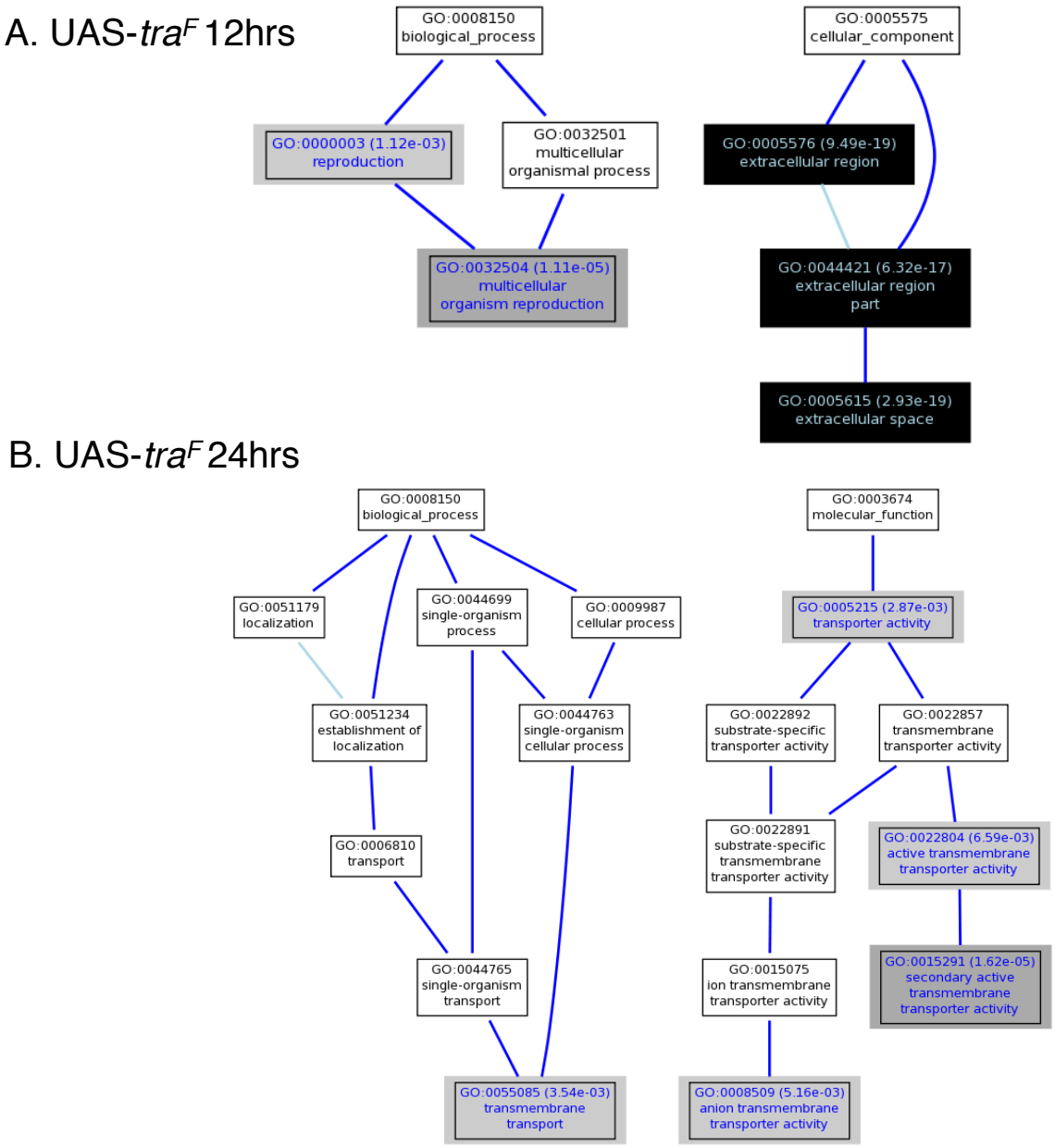


Figure 3.17. Graphical view of hierarchical terms relating to cellular components, biological and molecular processes enriched in UAS-*tra*^{F} fat body 12hrs following temperature shift. Terms are related to each other in a semi-hierarchical fashion (a directed graph structure), from very broad terms (at the top of the hierarchy) to specific (at the bottom).

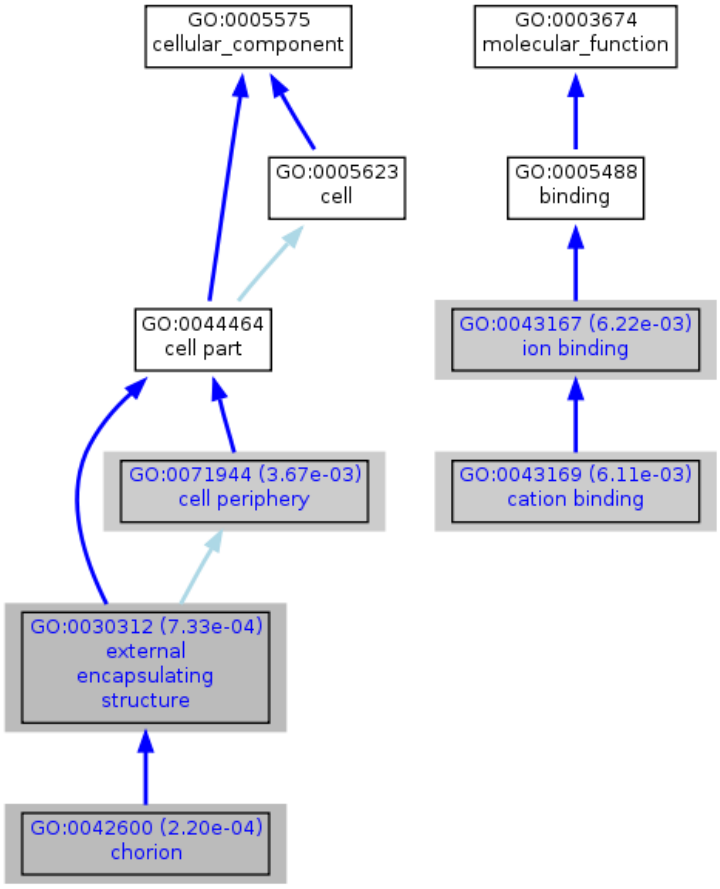


Figure 3.18. Graphical view of hierarchical terms relating to cellular components and molecular processes enriched in *tra2*^{ts} gonads 12hrs following temperature shift. Terms are related to each other in a semi-hierarchical fashion (a directed graph structure), from very broad terms (at the top of the hierarchy) to specific (at the bottom). No GO terms were enriched in the *tra2*^{ts} gonad 24hrs experiment.

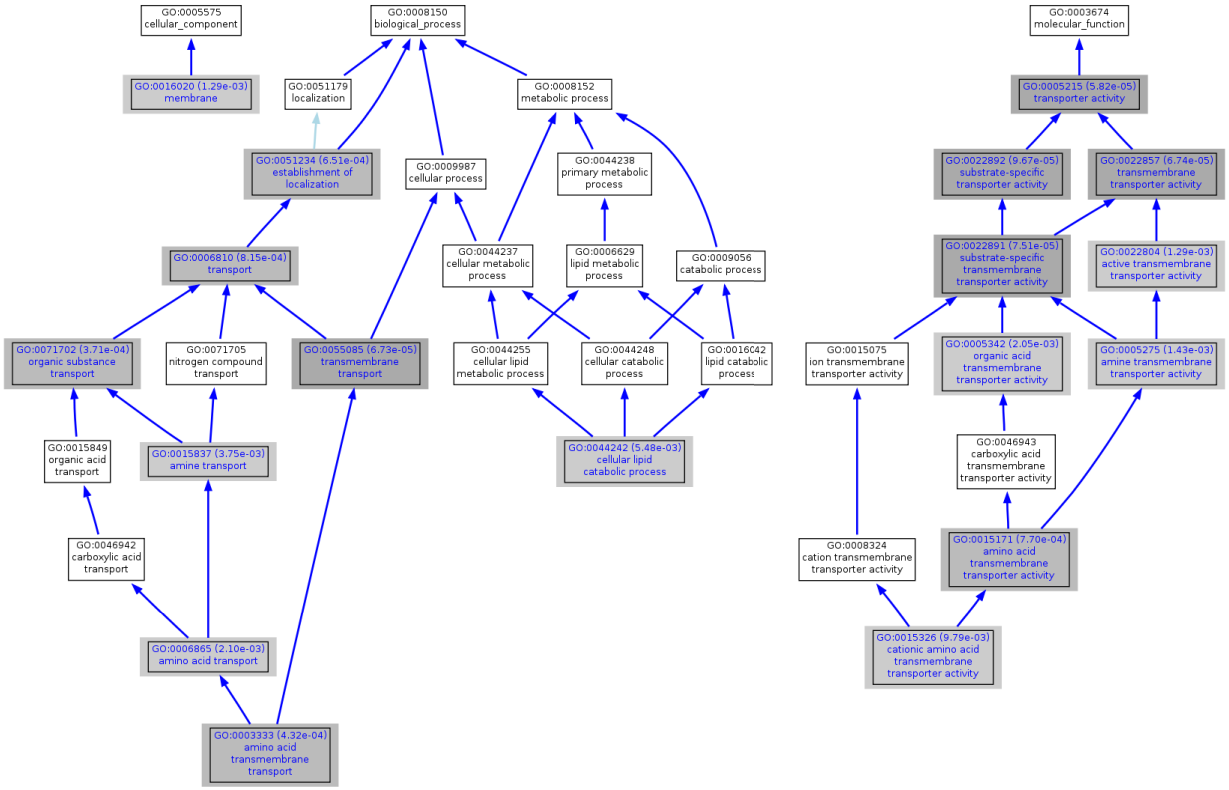


Figure 3.19. Graphical view of hierarchical terms relating to cellular components, biological and molecular processes enriched in UAS-*tra*^{F} gonads 12hrs following temperature shift. Terms are related to each other in a semi-hierarchical fashion (a directed graph structure), from very broad terms (at the top of the hierarchy) to specific (at the bottom). No GO terms were enriched in the UAS-*tra*^{F} gonad 24hrs experiment.

Occupancy and expression

To empirically determine where DSX binds in *D. melanogaster*, we collaborated with scientists at the National Institutes of Health (Brian Oliver, NIDDK), who performed chromatin immunoprecipitation followed by sequencing (ChIP-Seq) on S2 cells expressing tagged DSX^M or DSX^F, DSX^M or DSX^F DNA adenine methyltransferase identification (DamID) on ovary, and adult female and male fat body, followed by either sequencing (DamID-Seq) or hybridization to microarrays (DamID-chip) was also performed. These experiments allowed us to develop a list of candidate DSX targets while also providing insight on the context-specificity of DSX action. Our analyses identify 10411 genes with putative binding sites for DSX. A goal of this project was to identify biologically significant targets for DSX. To do this, we intersected these occupancy datasets with our *dsx*-dependent gene expression data to robustly determine direct DSX targets. We found that genes bound by DSX according to DamID-seq and –array datasets also changed in response to a switch in DSX isoforms in the fat body and gonad RNA-seq datasets, suggesting that identified targets are biologically meaningful (Table 3.15-3.19).

In order to estimate the probability for the same set of genes to appear in two top ranking gene lists between the two studies (occupancy and RNA-seq), we performed a hypergeometric distribution and Fisher's Exact test. The hypergeometric test was used to estimate the probability of the same set of genes (genes with occupancy) to appear in two top ranking lists in the two different studies (Occupancy vs RNA-seq). In parallel, the Fisher's exact test was used to examine the overlap between occupied and differentially expressed genes for experimental samples with controls used as null datasets. Datasets with significant enrichment of occupied genes in experimental samples relative to the controls (null datasets) had to possess a p-value<0.001. Expression datasets for each experiment and control were tested as indicated. We

expected that intersection of DSX occupancy datasets with expression data from control animals would have a hypergeometric probability equal to 1, since genes changing in control treated animals are responding to a change in temperature and not DSX isoforms. However, we expected that intersection of DSX occupancy datasets with expression data from DSX isoform switched animals to have a hypergeometric probability equal to 0, since genes changing in mutant animals are responding to a change in DSX isoform. We intersected our female fat body DamID-seq dataset with the female fat body RNA-seq datasets (Table 3.15-3.19). Intersection of datasets from control animals (w^{1118}) at 12 hr and 24 hr of treatment at 29°C resulted in a hypergeometric probability of ~1. This indicates that the genes differentially expressed in the fat body from control treated animals and the genes with DSX occupancy do not overlap significantly. However, in fat body from animals undergoing a change in DSX isoforms from DSX^F to DSX^M (XX $tra2^{ts}$) for 12 hrs and 24 hrs at 29°C, the hypergeometric probability is ~1 and ~0, respectively. This indicates that the genes in fat body of female animals undergoing a change in DSX isoform (DSX^F to DSX^M) overlap significantly with genes occupied by DSX only in the 24 hr time point. Along this line of investigation, we performed statistical tests to robustly determine targets of DSX for the remaining RNA-seq datasets of the male fat body and the gonads (Table 3.15-3.19). Interestingly, genes differentially expressed in ovaries of animals undergoing a change in DSX isoform from DSX^F to DSX^M (XX $tra2^{ts}$) was poorly enriched for DSX binding sites relative to controls using Hypergeometric distribution tests. The ovary contains heterogeneous cell types (germ cells and somatic cells) and not all somatic cells express DSX. Since a fraction of the soma expresses DSX, a possibility for the negative outcome of these experiments is that we are simply unable to capture genes expressed as a consequence of the change in DSX isoform due to the high contribution of RNA from the germline and other

somatic cell populations. Despite the fact that DSX binds to thousands of sites in the genome, astonishingly hundreds of genes are actually regulated by DSX in the fat body and gonads based on our RNA-seq analysis. The other occupied genes may well be regulated by DSX in the fat body or gonad under different conditions (e.g. nutritional response), at different times in development, or may be DSX targets in other tissues. We conclude that specificity of DSX action comes not from where DSX binds in the genome, but from its ability to coordinate with other tissue-specific and condition-specific transcription factors that dictate which possible DSX targets will actually be regulated at any one time or place.

As described earlier, the *yolk protein* genes (*Yp1/2*) are direct targets of DSX^F in female fat body tissue. These genes are activated by DSX^F and repressed by DSX^M and are therefore not expressed in males. To examine this model, where DSX isoforms have opposite regulatory functions in the two sexes, we compared gene expression of putative chromosomal targets in fat body from females and males undergoing a change in DSX isoform at 12 hrs and 24 hrs of the switch. In fat body, we found that genes occupied by DSX in $tra2^{ts}$ females progressively increased in repression upon the switch from DSX^F to DSX^M while genes activated in response to the DSX isoform switch decreased. The opposite result was found in UAS-tra^F male fat body. In UAS-tra^F male fat body, genes occupied by DSX progressively increased in activation upon the switch from DSX^M to DSX^F while genes repressed in response to the DSX isoform switch decreased. When we analyzed the mode of DSX regulation of occupied genes in female and male gonads, the directionality was consistent with the fat body. Similar to $tra2^{ts}$ female fat body, genes occupied by DSX in ovaries undergoing a switch from DSX^F to DSX^M were repressed. Though, the repression did not progressively increase over time, the number of genes repressed was consistent over time. In UAS-tra^F male testes, we also observed that genes were initially

repressed in response to the switch in DSX isoform from DSX^M to DSX^F 12 hrs of the temperature shift. However, by 24 hrs of the temperature shift from DSX^M to DSX^F, we found that occupied genes were activated in response to the DSX switch. Thus, genes that are occupied by DSX and regulated as a consequence of *dsx* are regulated in the *Yp*-like mode of regulation. Results are summarized in Table 3.19.

Discussion

We identified hundreds of genes regulated downstream of DSX isoform activity. DSX^M and DSX^F have differences in gene sets activated or repressed upon induction of a switch in DSX isoform usage in all tissues examined from $tra2^{ts}$ females and UAS-tra^F males, demonstrating that each DSX isoform has distinct biochemical activities (Figure 3.12, Table 3.8). We find that there are genes differentially expressed and shared between time points of a given genotype as well as differentially expressed genes unique to one time point for a given genotype (Figure 3.13). Differences observed in the identity of genes expressed in these experiments may be due to the differences in stoichiometric ratios of DSX^M and DSX^F induced after the switch in DSX isoform in $tra2^{ts}$ females and UAS-tra^F males. There may also be differences in stoichiometric ratios of DSX^M and DSX^F with other endogenous binding partners to confer regulation of direct and transcriptional targets. In context of our isoform switch experiments, DSX^M and DSX^F have different activities with respect to the genes that are induced or repressed, with many more genes having induced rather than repressed expression in females undergoing a switch from DSX^M and DSX^F and vice versa.

Despite the fact that DSX binds to thousands of sites in the genome, astonishingly hundreds of genes are actually regulated by DSX in the fat body and gonads based on our RNA-

seq analysis. Statistical analyses comparing individual occupancy and expression datasets showed a significant enrichment of DSX target genes in our expression datasets (Table 3.9-3.12). Although DSX occupies the same genes in multiple tissues and DSX^M/DSX^F isoforms have similar binding patterns, our results suggest that many genes are poised to respond to DSX, but that there are additional stimuli (temporal/spatial, hormonal/nutritional, or sex-specific factors) that influence DSX isoform activity since induction of either DSX^M or DSX^F in the opposite sexes induces or repressed different genes. Another possibility is that while known DSX targets, such as the yolk proteins, are constitutively regulated in the adult, there is a subset of genes bound by DSX and also transcriptionally regulated throughout adulthood. Perhaps putative direct targets of DSX are regulated during development or in another context, but are *dsx*-independent in the adult fat body. From our intersection between our occupancy and expression profiling datasets, we conclude that DSX regulatory specificity depends on where DSX is bound and the ability to coordinate with sex-, tissue-, or condition-specific factors that dictate which possible DSX targets will actually be regulated at any one time or place.

Table 3.15-3.19. Intersection between occupancy and expression datasets. The hypergeometric test was used to estimate the probability of the same set of genes (genes with occupancy) to appear in two top ranking lists in two different studies (Occupancy vs RNA-seq). In parallel, the Fisher's exact test was used to examine the overlap between occupied and differentially expressed genes for experimental samples with controls used as null datasets. Datasets with significant enrichment of occupied genes in experimental samples relative to the controls (null datasets) had to possess a p-value<0.001. Expression datasets for each experiment and control were tested as indicated.

Table 3.15. Intersection between occupancy (female fat body DamID-seq and -array, male fat body DamID-seq) and fat body expression datasets. Datasets with significant enrichment in occupancy in yellow.

Fat body Sample Name	female fat body DamID-seq (6072 genes occupied)		Fisher's Exact Test	Hypergeometric Test
	with occupancy	without occupancy		
Female fat body tra2ts12hr	87	211	0.06658	0.9997802
Female fat body tra2ts24hr	269	205	3.36E-04	4.77E-14
Female fat body wt12hr	3	19		0.999598
Female fat body wt24hr	1	17		1
Male fat body traF12hr	32	72	0.02071	0.9794391
Male fat body traF24hr	49	42	0.05195	0.006584294
Male fat body control12hr	70	93		0.07809571
Male fat body control24hr	43	70		0.5980417
	female fat body DamID-array (5838 genes occupied)		Fisher's Exact Test	Hypergeometric Test
	with occupancy	without occupancy		
Female fat body tra2ts12hr	66	232	3.14E-01	1
Female fat body tra2ts24hr	204	270	0.001005	0.006197699
Female fat body wt12hr	3	22		0.9968017
Female fat body wt24hr	1	17		0.9997846
Male fat body traF12hr	28	76	0.08443	0.9911026
Male fat body traF24hr	33	58	0.5529	0.6289365
Male fat body control12hr	61	102		0.5294772
Male fat body control24hr	36	77		0.9084043
	male fat body DamID-seq (5264 genes occupied)		Fisher's Exact Test	Hypergeometric Test
	with occupancy	without occupancy		
Female fat body tra2ts12hr	63	235	2.71E-01	0.9999995
Female fat body tra2ts24hr	208	266	4.33E-05	2.04E-06
Female fat body wt12hr	2	20		0.9985831
Female fat body wt24hr	0	18		1
Male fat body traF12hr	23	81	0.0286	0.9965996
Male fat body traF24hr	38	53	0.1872	0.06703983
Male fat body control12hr	57	106		0.3981217
Male fat body control24hr	36	77		0.6972963

Table 3.16. Intersection between occupancy (ovary DamID-seq, S2 ChIP-seq) and fat body expression datasets. Datasets with significant enrichment of occupied genes in yellow.

Fat body Sample Name	female ovary DamID-seq (5821 genes occupied)		Fisher's Exact Test	Hypergeometric Test
	with occupancy	without occupancy		
Female fat body tra2ts12hr	94	204	0.09386	0.9846005
Female fat body tra2ts24hr	220	254	2.65E-05	2.41E-05
Female fat body wt12hr	3	19		0.9967052
Female fat body wt24hr	0	18		1
Male fat body traF12hr	27	77	0.1077	0.9948116
Male fat body traF24hr	31	60	1	0.7723122
Male fat body control12hr	58	105		0.7037468
Male fat body control24hr	39	74		0.7613577
	DSX^M S2 ChIP-seq (5258 genes occupied)		Fisher's Exact Test	Hypergeometric Test
	with occupancy	without occupancy		
Female fat body tra2ts12hr	99	199	0.3561	0.5922105
Female fat body tra2ts24hr	192	282	0.01268	0.000985738
Female fat body wt12hr	5	17		0.9099446
Female fat body wt24hr	2	16		0.9938076
Male fat body traF12hr	17	87	0.001169	0.9999793
Male fat body traF24hr	27	64	0.7625	0.8227034
Male fat body control12hr	57	106		0.3940826
Male fat body control24hr	36	77		0.6942703
	DSX^F S2 ChIP-seq (6644 genes occupied)			
	with occupancy	without occupancy	Fisher's Exact Test	Hypergeometric Test
Female fat body tra2ts12hr	112	184	0.3689	0.9664152
Female fat body tra2ts24hr	223	251	0.05208	0.02608895
Female fat body wt12hr	6	16		0.9556815
Female fat body wt24hr	4	14		0.9802283
Male fat body traF12hr	23	81	0.002149	0.9999967
Male fat body traF24hr	33	58	1	0.9090355
Male fat body control12hr	66	97		0.7327475
Male fat body control24hr	40	73		0.9512268

Table 3.17. Intersection between occupancy (female fat body DamID-seq and -array, male fat body DamID-seq) and gonad expression datasets. Datasets with significant enrichment of occupied genes in yellow.

Gonad Sample Name	female fat body DamID-seq (6072 genes occupied)		Fisher's Exact Test	Hypergeometric Test
	with occupancy	without occupancy		
Female ovary tra2ts12hr	99	72	0.7269	3.516997E-07
Female ovary tra2ts24hr	9	38	0.004386	0.9989643
Female ovary wt12hr	74	59		6.501467E-05
Female ovary wt24hr	38	48		0.1854159
Male testis traF12hr	54	59	0.000126	0.03361305
Male testis traF24hr	49	60	0.1418	0.1159789
Male testis control12hr	24	83		0.9999102
Male testis control24hr	69	122		0.8079622
	female fat body DamID-array (5838 genes occupied)		Fisher's Exact Test	Hypergeometric Test
	with occupancy	without occupancy		
Female ovary tra2ts12hr	75	96	0.5631	0.04843507
Female ovary tra2ts24hr	11	36	0.08396	0.9862914
Female ovary wt12hr	63	70		0.01171685
Female ovary wt24hr	34	52		0.381026
Male testis traF12hr	46	67	1.008E-05	0.2637482
Male testis traF24hr	41	68	0.6176	0.5195572
Male testis control12hr	15	92		1
Male testis control24hr	66	125		0.8152643
	male fat body DamID-seq (5264 genes occupied)			
	with occupancy	without occupancy	Fisher's Exact Test	Hypergeometric Test
Female ovary tra2ts12hr	74	97	0.4181	0.005697061
Female ovary tra2ts24hr	6	41	0.007506	0.9997388
Female ovary wt12hr	64	69		0.0004010984
Female ovary wt24hr	30	56		0.4513026
Male testis traF12hr	45	68	1.836E-05	0.5236897
Male testis traF24hr	38	71	0.5218	0.4375067
Male testis control12hr	15	92		0.9999991
Male testis control24hr	59	132		0.819703

Table 3.18. Intersection between occupancy (ovary DamID-seq, S2 ChIP-seq) and gonad expression datasets. Datasets with significant enrichment of occupied genes in yellow.

Gonad Sample Name	male fat body DamID-seq (5264 genes occupied)		Fisher's Exact Test	Hypergeometric Test
	with occupancy	without occupancy		
Female ovary tra2ts12hr	88	83	0.105	0.0001048439
Female ovary tra2ts24hr	9	38	4.30E-06	9.98E-01
Female ovary wt12hr	81	52		2.62E-08
Female ovary wt24hr	52	34		1.06E-05
Male testis traF12hr	46	67	0.0002096	0.2558999
Male testis traF24hr	46	63	0.004777	0.1683597
Male testis control12hr	19	88		0.9999972
Male testis control24hr	50	141		0.9996053
	DSX^M S2 ChIP-seq (5258 genes occupied)		Fisher's Exact Test	Hypergeometric Test
	with occupancy	without occupancy		
Female ovary tra2ts12hr	79	92	0.2034	0.1884977
Female ovary tra2ts24hr	8	39	0.002089	0.999954
Female ovary wt12hr	72	61		0.004640164
Female ovary wt24hr	38	48		0.422159
Male testis traF12hr	44	69	0.0005797	0.8110039
Male testis traF24hr	44	65	0.6222	0.7136388
Male testis control12hr	19	88		1
Male testis control24hr	71	120		0.9455234
	DSX^F S2 ChIP-seq (6644 genes occupied)			
	with occupancy	without occupancy	Fisher's Exact Test	Hypergeometric Test
Female ovary tra2ts12hr	74	97	0.2026	0.005522157
Female ovary tra2ts24hr	5	42	7.70E-05	0.9999427
Female ovary wt12hr	68	65		2.40E-05
Female ovary wt24hr	38	48		0.02740082
Male testis traF12hr	39	74	2.062E-05	0.4627068
Male testis traF24hr	39	70	0.3052	0.3564657
Male testis control12hr	11	96		1
Male testis control24hr	57	134		0.8881713

Table 3.19. Mode of DSX regulation on occupied genes. Numbers and percentages of activated and repressed genes having occupancy in each experiment.

	Genes with DSX occupancy				
Sample Name	Repressed genes	Activated genes	Total # of differential genes	% Repressed genes	% Activated genes
Female fat body tra2^{ts}12hr	45	41	87	52	47
Female fat body tra2ts24hr	189	75	264	72	28
Male fat body traF12hr	14	17	31	45	55
Male fat body traF24hr	14	34	48	29	71
Ovary tra2ts12hr	75	24	99	76	24
Ovary tra2ts24hr	7	2	9	78	22
Testis traF12hr	38	8	46	83	17
Testis traF24hr	21	28	49	43	57

CHAPTER 4: DOUBLESEX REGULATION OF ECDYSONE SIGNALING IN THE GONAD

Summary

The current thinking is that sex-specific steroid hormone signaling is not present in insects (or other non-vertebrates) and that the invertebrate steroid hormone ecdysone is used similarly in both sexes as an important regulator of developmental timing. In this chapter we explore the role of sex-specific steroid hormone signaling in the *Drosophila* gonad and identify the ecdysone receptor (EcR) as a key target gene which *dsx* regulates to control properly dimorphic niche formation. We find that ecdysone signaling has sex-specific action in *Drosophila*, which completely changes our view of such hormones as being vertebrate adaptations for creating sexual dimorphism. These experiments make Drosophila an invaluable model for understanding how such hormones control sex characteristics and shed light on how steroid hormones evolved from being nutritional signals (derived from cholesterol) to sex hormones.

Introduction

Diverse sex-determination mechanisms are utilized by species throughout the animal kingdom to create sexual dimorphism in the gonad. In mammals, sex determination depends on the presence or absence of expression of the transcription factor sex-determining region on the Y (Sry) located on the male-specific Y chromosome (Gubbay et al., 1990; Sinclair et al., 1990). Expression of Sry in the genital ridges (gonadal primordium) of the biopotential gonad results in their development of the testes and repression of the female sex-determining pathway, whereas in the complete absence of Sry, ovary development occurs (Koopman et al., 1991). The end result of Sry expression in the genital ridges, are sex-specific gonads that secrete gonadal hormones (estrogen in females and testosterone in males) that initialize development of secondary sex characteristics specific to the male or female developmental programs.

Like mammals, sex differentiation is generally thought to be determined on a genetic basis in *Drosophila* and converge on terminal differentiation genes, such as *doublesex*/DMRTs, that mediate development of sexually dimorphic gonads and male vs. female characteristics. Unlike vertebrates which have sex-specific steroids that play a pivotal role in sex differentiation (Nakamura et al., 2010), the traditional view is that sex-specific steroid hormone signaling does not exist in Drosophila (Gilbert, 2006).

During development, invertebrate steroid signaling is thought to be used similarly in both sexes as an important regulator of developmental timing rather than to direct sexual fate (Negri and Pellecchia, 2012). However, our preliminary data challenge this idea and indicate that sex-specific steroid hormone signaling is present during gonad development in *Drosophila* and is critical for the proper sexual development of the male vs. female gonad stem cell niches. Our data indicate that, rather than being regulated at the level of systemic hormone levels, as is seen

in mammals, steroid hormone signaling in Drosophila is regulated by local hormone response, modulated by *dsx*, to control sexual dimorphism in the germline stem cell niche.

As discussed in Chapter 1, a critical aspect of gonad development is the formation of the male and female germline stem cell niches. In wild-type *Drosophila*, both the ovary and testis have germline stem cells that are controlled and maintained by surrounding somatic cells which form their niches. Critical components of these niches are the "hub" in males and "terminal filaments and cap cells in females (Hardy, 1979). In wild-type gonads, the hub develops in males at the end of embryogenesis, while the terminal filaments form in females during late larval/ pupal stages. In the absence of *dsx*, we find that all *dsx* mutant gonads appear fully male-like at the end of embryogenesis regardless of their chromosomal constitution. This is in contrast to to wild-type embryonic gonads, where males have a hub and the female gonad does not. In late larvae of *dsx* mutants, at the time when terminal filaments normally form, we find that in 50% of animals the hubs were lost and terminal filaments formed in their place. This occurred in 50% of both XY and XX animals. In both XX and XY adults, 50% of the gonads still had hubs and 50% had terminal filaments.

Since both XX and XY gonads form hubs in *dsx* mutants, *dsx* is required in females to block hub formation at the time that the hub normally forms in the embryo. The niches are more plastic than previously thought. Even though a hub forms in *dsx* mutants, it is still sensitive to feminizing influence and can change fate and form terminal filaments. In *dsx* mutant larvae, 50% of both XX and XY gonads switch from hub to terminal filaments. Thus, *dsx* is normally required in females to ensure a robust response to the "pro terminal filament" pathway so that all female gonads form terminal filaments. Further, *dsx* is normally required in males to repress the "pro terminal filament" pathway and ensure that all male gonads maintain hub fate.

Among the genes identified in our whole-genome analysis and RNAi screen (Chapters 2) is the *Ecdysone receptor (EcR)* gene involved in steroid hormone signaling. Work by others has shown that *EcR* is important for ovary morphogenesis (Chapter 1). Thus, it represents the best candidate for the "pro female gene" in our model of sex-specific gonad niche development. Even though other tissues in the male must respond to ecdysone to control metamorphosis, the testis remains much as it has since the end of embryogenesis with continued spermatogenesis. We identify *Ecdysone receptor* as a "pro female gene" and find that *dsx* allows the steroid hormone ecdysone to act differently in males and females, so that males ignore this signal and maintain the hub and testis, while females produce a robust response to form ovaries. Our data suggests the intriguing hypothesis that sex-specific steroid hormone signaling is important for sexually dimorphic gonad development in *Drosophila*, as it is in mammals.

Ecdysone receptor (*EcR*) encodes a nuclear receptor that binds to 20-hydroxyecdysone, the master regulatory hormone that directs metamorphosis. The *EcR* locus spans 77kb. Through the use of two promoters and alternative splicing, *EcR* encodes three isoforms (EcRA, EcRB1, EcRB2). All three EcR isoforms share the same DNA binding domain but have variable N-terminal domains (Koelle et al., 1991; Talbot et al., 1993). EcR can bind to ecdysone alone, but optimal binding of EcR to its targets requires its co-factor USP (Grad et al., 2001; Grebe et al., 2004). USP can interact with all three EcR isoforms to form DNA-binding heterodimers (Yao et al., 1992; Bender et al., 1997). In the presence of ecdysone, the 20E/EcR/USP heterodimer is stabilized and binds to DNA encoding a canonical ecdysone response elements (EcRE) which results in transcriptional activation of ecdysone responsive genes (Reviewed in Quinn et al., 2012). Genes directly activated by the 20E/EcR/USP complex include "primary response genes", *Broad-Complex* (BR-C), *E74*, and *E75*, which encode transcription factors that control

expression of a battery of genes associated with biological changes resulting from the ecdysone pulse (Figure 4.1).

In this chapter, we investigate the role of sex-specific steroid hormone signaling in the *Drosophila* gonad. We have found that males and females exhibit a dramatically different response to ecdysone in the gonads. This differential response is likely due to the ecdysone receptor (*EcR*) expressed only in the female and not the male gonad. Using a sensitized genetic background, dsx^{D}/+, we find that the sex-specific ecdysone response that we have observed in the gonad has a functional consequence for sex-specific development of the gonad stem cell niche. These results challenge the concept that sex-specific steroid hormone signaling is a universal regulator of sexual dimorphism in gonad development across the animal kingdom.

Materials and methods

Fly stocks

P{hs-GAL4-EcR.LBD}SBM, P{UAS-GFP.nls}14, P{EcRE.lacZ}SS4, Dp(1;Y)BS; dsx1 pp/TM3, Sb1, Df(3R) dsx^{3}/TM3, Sb^{1}, st1 βTub85DD ss1 es/TM2 & dsxD Sb1 e1/TM2 (Kyoto, DGRC), *P{UAS-EcR.C}TP1-4, P{TRiP.JF02538}attP2 (EcR RNAi), Traffic jam-GAL4, C587-GAL4, dsx-GAL4* (S. Goodwin), *Unpaired-GAL4, P{UAS-EcR.A.W650A}TP5, P{UAS-EcR.B1-ΔC655.W650A}TP1-9, P{GMR46B07-GAL4}attP2* (Line 46268, Janelia GAL4), *P{GMR44C04-GAL4}attP2* (Line 45719, Janelia GAL4), *P{GMR46E08-GAL4}attP2* (Line 48167, Janelia GAL4), *P{GMR44B03-GAL4}attP2* (Line 50200, Janelia GAL4) were obtained from the Bloomington Stock Center unless otherwise indicated.

Immunohistochemistry and all sample imaging

Tissue was dissected from larvae or adult flies aged 1 to 3 days in PBS followed by fixation for 10 minutes in PBS containing 4% formaldehyde and rinsed 2 X 10 minutes in PBTx. Samples were blocked in PBTx with 1% BSA for at least 1 hour and then incubated in BBTx with primary antibody overnight at 4 °C. Following 3 X 10 minute washes in BBTx, samples were incubated in BBTx with secondary antibody for 3-4 hours at room temperature. Following 3 X10 minute washes in PBTx. Staining was performed as described in (Gonczy, 1997), and samples were mounted on slides in 2.5% DABCO. The following primary antibodies were used: Chicken anti-VASA (R. Lehmann) at 1:10,000; rat anti-DCadherin2 (DSHB) at 1:20, rat anti-NCadherin (DSHB) at 1:20, rabbit anti-GFP (Torrey Pines) at 1:2,000, rabbit anti-betagalactosidase (Cappel) at 1:10,000, mouse anit-EcR (Ag10.2 EcR common DSHB) at 1:10, mouse anti-Broad (25E9.D7 Broad core DSHB) at 1:20, ginea pig anti-traffic jam (M. Van Doren) at 1:100. The following secondary antibodies were used: Alexa 546 goat anti-chicken at 1:500; Alexa 633 goat anti-chicken at 1:500; Alexa 488 goat anti-rat at 1:300-500; Alexa 633 goat anti-mouse at 1:500; Alexa 546 goat anti-mouse at 1:500; Alexa 488 goat anti-guinea pig at 1:500, Alexa 488 goat anti-rabbit at 1:500. All Alexa probes are from Molecular Probes (Invitrogen, Carlsbad, CA). All immunohistochemistry samples were imaged on a Zeiss LSM 700 confocal microscope.

Developmental staging and heat induction of GAL4-LBD fusion proteins

To obtain flies in similar developmental stages, *hs-GAL4-EcR.LBD; UAS-GFPnls* flies were transferred into a fresh vial to lay eggs for 4-6 hours, and were then removed. Vials were left at 25°C for 96 h (mid third instar, ML3). All heat treatments were performed by incubating

plastic culture vials containing food in a 37°C water bath for 30 minutes followed by 30 minutes of rest. In total, larvae were heat treated for 2.5 hours and then allowed to recover at 25°C for 4-6 hours.

For larval staging

To obtain flies in similar developmental stages, flies were transferred into a fresh vial to lay eggs for 4-6 hours, and were then removed. Vials were left at 25° for 96 h (mid third instar, ML3) or 120 h (late third instar, LL3).

20-Hydroxyecdysone feeding

For ecdysone feeding, a 1mg/ml 20-hydroxyecdysone (Sigma) stock was diluted with 60% ethanol and added to dry yeast until a paste was made. Control contained water only. To obtain flies in similar developmental stages, flies were transferred into a fresh cage to lay eggs for 4-6 hours on apple juice plates, and were then removed. Yeast paste with 20-hydroxyecdysone or water (for control) were added to apple juice plates with embryos and were left at 25° for 96 h (mid third instar, ML3) or 120 h (late third instar, LL3). Fresh yeast paste made from 20-hydroxyecdysone solution or water (control) was added to apple juice plates daily until the day of dissection.

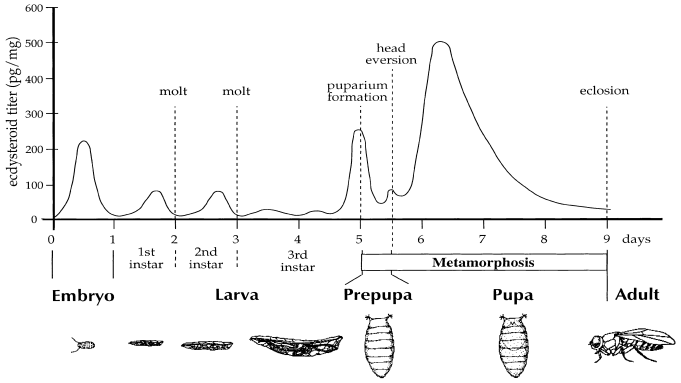
Ecdysone reporter construct

To generate Construct 1, primers flanked by restriction sites (BamHI and BglII) were used to amplify a 3.5kb genomic region from a genomic clone, BACR08A11 (BACPAC

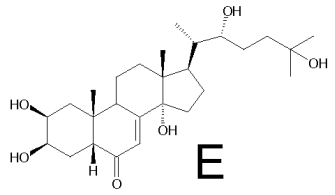
Resources Center). The PCR product was cloned into the vector PJR20 (R. Johnston) between the sites BamHI and BglII, resulting in a 11kb Construct 1.

Figure 4.1. The ecdysone signaling pathway. A) Ecdysone pulses from the prothoracic gland are required for all aspects of morphogenesis including the formation of the body plan during late embryogenesis, hatching and development of the first larval instar, and first and second instar cuticle molting. Finally, the prothoracic gland secretes a large ecdysone pulse at the end of third instar responsible for the larval to pupal transition which marks the beginning of adult tissue metamorphosis. B) During larval stages, ecdysone is produced in the prothoracic gland which is an endocrine tissue that expresses genes encoding ecdysone biosynthetic enzymes and is responsible for generating the ecdysone pulse. Once ecdysone is released into the hemolymph, ecdysone is converted into an active form, 20-hydroxyecdysone (20E), by Shade, a P450 monooxygenase that is expressed in non endocrine tissues. C) 20-hydroxyecdysone binds and activates the nuclear receptor Ecdysone receptor (EcR) and its receptor binding partner Ultraspiracle (USP). EcR can bind to ecdysone alone, but optimal binding of EcR to its targets requires its co-factor USP. In the presence of ecdysone, the 20E/EcR/USP heterodimer is stabilized and binds to DNA encoding a canonical ecdysone response elements (EcRE) which results in transcriptional activation of ecdysone responsive genes. D) Genes directly activated by the 20E/EcR/USP complex include “primary response genes”, Broad-Complex (BR-C), Eip74, and Eip75, which encode transcription factors that control expression of a battery of genes associated with biological changes resulting from the ecdysone pulse (Figure 4.1).

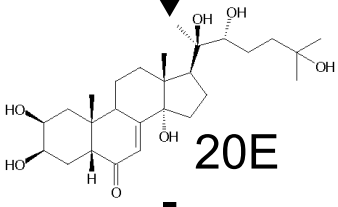
A. Ecdysone pulses regulate metamorphosis.



B. Synthesis

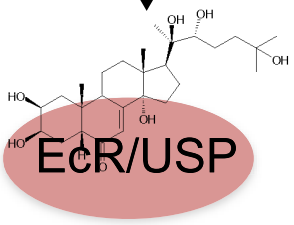


E



20E

C. Reception



D. Co-regulation

Early genes

BR-C, Eip74, Eip75, Eip93, crol, Hr78

Results

Spatial activity patterns for EcR in the larval gonad

In contrasts to our understanding of the transcriptional responses to ecdysone signaling in the gonad, temporal and spatial specificity of ecdysone signaling is poorly understood. Ecdysone hormone binds and activates the nuclear receptor, ecdysone receptor (EcR), and its binding partner, Ultraspiracle (USP) (Thummel 1996, 1990, 1995; Koelle et al. 1991). EcR can bind to the ecdysone hormone alone. However, in order for EcR to bind to ecdysone responsive elements and activate transcription of its targets, EcR requires a co-factor USP. In the presence of ecdysone hormone, EcR dimerizes to USP and the EcR/USP/20E complex binds to ecdysone responsive elements to transcriptionally activate genes. In the absence of ecdysone hormone, the presence of EcR results in repression of target genes. Thus, differential responses to the ecdysone pulse during development can be achieved by restricting the biological response through the tissue-specific patterns of EcR isoform expression (Talbot et al., 1993; Truman et al., 1994).

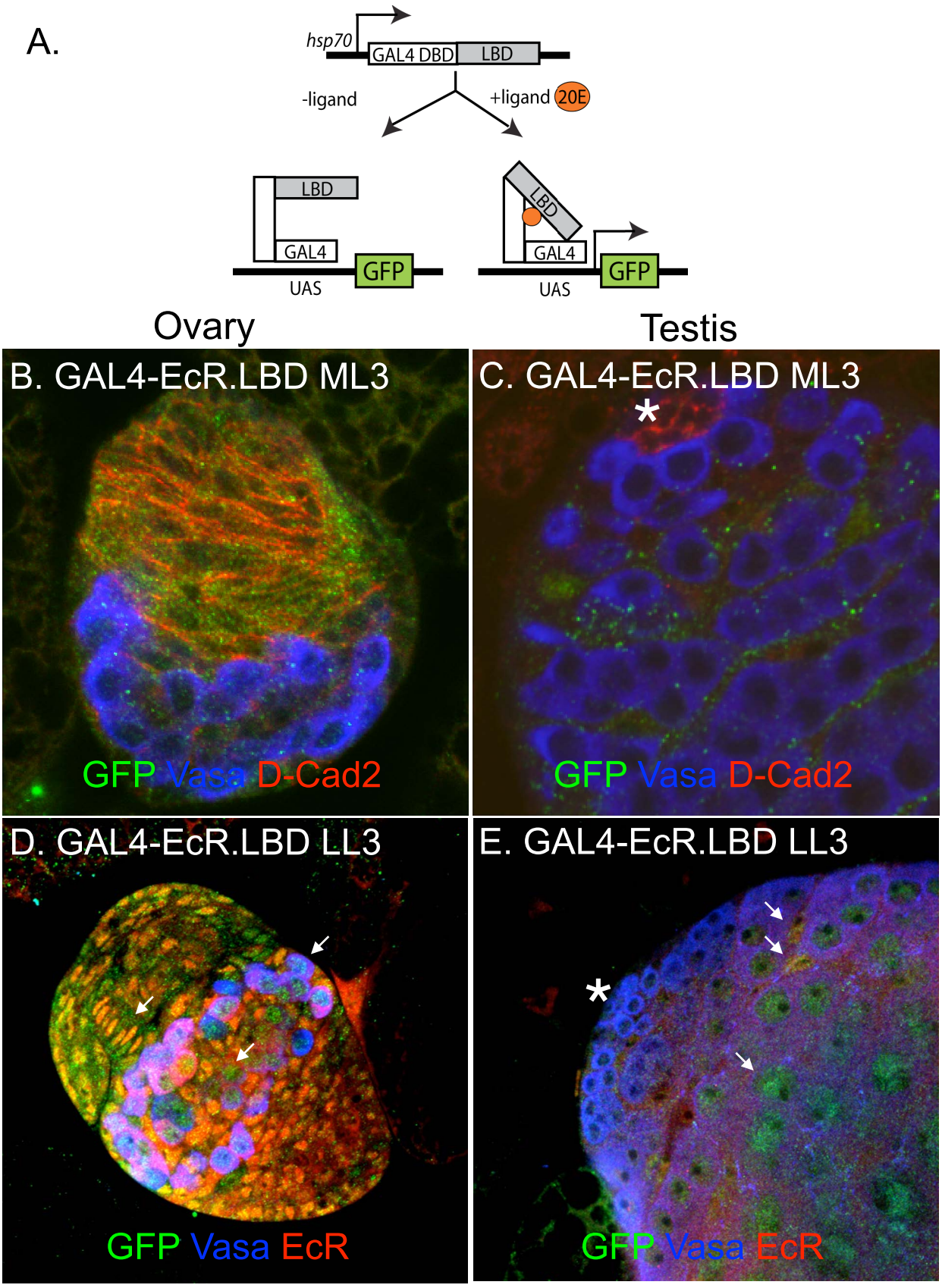
To determine where the ecdysone nuclear receptor, EcR, may be regulated by its ligand hormone, ecdysone, we utilized a 'ligand sensor' system to visualize spatial activity patterns for EcR in the larval gonad (Palanker et al., 2006; Kozlova and Thummel, 2002). This system consists of the Ecdysone Receptor ligand binding domain (LBD) tethered to the DNA-binding domain of the yeast GAL4 transcriptional activator. UAS-controlled GFP is used as the reporter. When this hybrid protein is expressed ubiquitously using the *hsp70* heat shock promoter, it activates GFP preferentially in regions of high ecdysone titer and co-activator protein USP (Figure 4.2A). The temporal and spatial pattern of the reporter expression indicates where and when the EcR ligand binding domain has been activated by its ligand, ecdysone. This experiment

provides a more direct means to follow the pattern of ecdysone hormone signaling in the context of the developing gonad (Kozlova and Thummel, 2002).

We examined the spatial patterns of ligand sensor activation in male and female gonads in mid third instar larvae (~96 hrs, After Egg Laying "AEL") when ecdysteroid titers are low and in late third instar larvae (~120 hrs AEL) when ecdysteroid titers are high at the onset of metamorphosis. In mid third instar (ML3) larval gonads, we detected activation of the EcR ligand sensor in the somatic cells of the ovary at higher levels compared to the testis (Figure 4.2BC). In the ML3 ovary, activation was displayed in the anterior soma and in the intermingled cells (Figure 4.2B), whereas in the ML3 testis, we observe activation of the ligand sensor in early somatic cells surrounding the niche (Figure 4.2C).

In contrast, a different pattern emerged when we analyzed ligand sensor activity at the onset of metamorphosis, ~120 hrs AEL, when ecdysteroid titers were high. In comparison to activation we observed at ML3, when ecdysteroid titers are low, we observed high activity of the ligand sensor in LL3 which is consistent with high ecdysteroid titer at this time. Similar to what we observed during ML3, we detected activation of the EcR ligand sensor in the ovary at higher levels compared to the testis (Figure 4.2DE). In the LL3 ovary, activation was displayed in the terminal filaments, a component of the female germline stem cell niche. Activation was also detected throughout the anterior soma, in intermingled cells and in germ cells. With the exception of ligand sensor activation in the germ cells of the ovary, GFP reporter expression co-localizes to the endogenous pattern of EcR detected by antibody staining (Figure 4.2D). In the testis, the ligand sensor did not display detectable activation in the hub, a component of the male germline stem cell niche. Activation of the ligand sensor was detected in some early somatic

Figure 4.2. Sex specific patterns of ecdysteroid receptor activation at the onset of metamorphosis. A) Schematic of paradigm. Mid third instar (ML3, ~96 hrs AEL) and late third instar (LL3, ~120 hrs AEL) larvae were heat shocked at 37°C for 2.5 hours and allowed to recover at 25°C for 4-6 hours. B) Female larval ovaries (ML3) *hs EcR LBD>GFP*. C) Male larval testes (ML3) hs EcR LBD>GFP. D) Female larval ovaries (LL3) *hs EcR LBD>GFP*. E) Male larval testes (LL3) *hs EcR LBD>GFP*.


A.
hsp70
GAL4 DBD
LBD
-ligand
+ligand
20E
LBD
GAL4
GFP
UAS
LBD
GAL4
GFP
UAS
Ovary
Testis
B. GAL4-EcR.LBD ML3
GFP Vasa D-Cad2
C. GAL4-EcR.LBD ML3
*
GFP Vasa D-Cad2
D. GAL4-EcR.LBD LL3
GFP Vasa EcR
E. GAL4-EcR.LBD LL3
*
GFP Vasa EcR

cells, in late somatic cells, and in germ cells (Figure 4.2E). GFP reporter expression present in early somatic cells co-labels with EcR detected by antibody staining.

When we examined activity of the ecdysone ligand sensor in *dsx* mutant gonads with a male niche component, the hub, we detect activation throughout the soma of all gonads with hubs as well as in germ cells. Activation was displayed in the cells of the hub and in the terminal epithelium (Figure 4.5D). This data suggests the possibility that *dsx* influences local ecdysone titer in gonads.

Ecdysone signaling response in female vs. male gonads

We examined the ecdysone signaling response in gonads using three different methods; a reporter to monitor EcR activation, direct antibody staining of EcR, and activation of a primary response gene, *Broad-C*, induced by the ecdysone signaling pathway.

To evaluate the ecdysone signaling response in male and female gonads, we first examined the spatial pattern of a reporter for ecdysone receptor activity (ecdysone receptor DNA binding sites upstream of a lacZ reporter, EcRE-lacZ). In the presence of ecdysone hormone, the 20E/EcR/USP heterodimer is stabilized and binds to DNA encoding ecdysone response elements (EcRE) which results in transcription of ecdysone responsive genes (Reviewed in Quinn et al., 2012). The EcRE-lacZ reporter is a transgene containing seven copies of an EcR responsive element (EcRE) upstream of a minimal promoter (hsp70) and the E. coli lacZ gene. The transgene is present in every cell, but only cells that contain ecdysone hormone and the nuclear receptor, EcR, and it cofactors are competent to respond to it (Kozlova T, Thummel CS) (Figure 4.3A).

We find that the larval ovary shows robust activity of this reporter while the same stage larval testis shows little activity (Figure 4.3AB). In the ovary, we see robust expression of the reporter in the terminal filaments, anterior soma, in intermingled cells and in basal cells. No signal was present in the germ cells of the larval ovary (Figure 4.3B). In the testis, the late somatic cells display weak expression of the reporter. Only one region of the testis (the terminal epithelium), which is at the opposite end of the testis from the hub, shows robust EcRE-lacZ staining (Figure 4.3C).

We also examined the spatial pattern of EcRE-lacZ activity in *dsx* mutant gonads. Due to the nature of the combination of *dsx* alleles and the EcRE-lacZ reporter, no gonads with hubs or terminal filaments were observed. Instead, gonads with no male-like or female-like niches were observed in this genetic background. In these "niche-less" *dsx* mutant gonads carrying the EcRE-lacZ reporter, we detect activation in the anterior and posterior soma of gonads (Figure 4.5C). Although additional experiments are required to determine that *dsx* is instructive in ensuring ecdysone signaling is activated in the correct genotype (XX vs. XY), this data suggests that *dsx* is not required for ecdysone signaling.

To understand if the differences we observed in the ecdysone signaling response in female and male gonads was due to differential expression of the Ecdysone receptor gene, we examined the expression of EcR in larval gonads from males and females (Figure 4.4AB). An antibody directed against the common region of all three EcR isoforms (EcRA, EcRB1, EcRB2) stained all somatic nuclei of the larval ovary. No EcR staining was observed within the PGCs (Figure 4.4A). This finding is in agreement with somatic expression of EcR previously reported in (Hodin and Riddiford 1998). When we examined larval male gonads of the same stage, we do not detect EcR staining in the somatic cells of the testis at comparable levels to the larval ovary.

Few early and late somatic cells are weakly labeled. Only pigment cells surrounding the testis and one region of the testis (the terminal epithelium), which is at the opposite end from the hub, shows strong somatic nuclear staining (Figure 4.4B).

While the Broad-Complex is not entirely dependent on ecdysone signaling, we examined expression of the transcription factor *Broad* (*br*), a downstream target of the pathway (Brennan 2001, Gancz et al., 2011). There are four different transcripts encoded by the broad locus: broad-Z1, Z2, Z3, and Z4 (DiBello). When we examined expression using an antibody directed against the common region of all Broad isoforms, we observed staining in all somatic nuclei in the larval ovary, but not the testis (Figure 4.4CD). In the testis, broad expression is restricted to a subset of early somatic cells of the testis. Similar to EcR localization in the larval testis, we also find strong somatic nuclear staining in the terminal epithelium (Figure 4.4D). Neither male nor female gonads have Broad expression in germ cells. These results are in agreement with somatic expression of Broad previously reported (Gancz et al., 2011; Mugat et al., 2000).

All together, these results demonstrate that males and females exhibit a dramatically different response to ecdysone in the gonads. This difference may be due to sexually dimorphic EcR expression in the larval ovary vs. the testis.

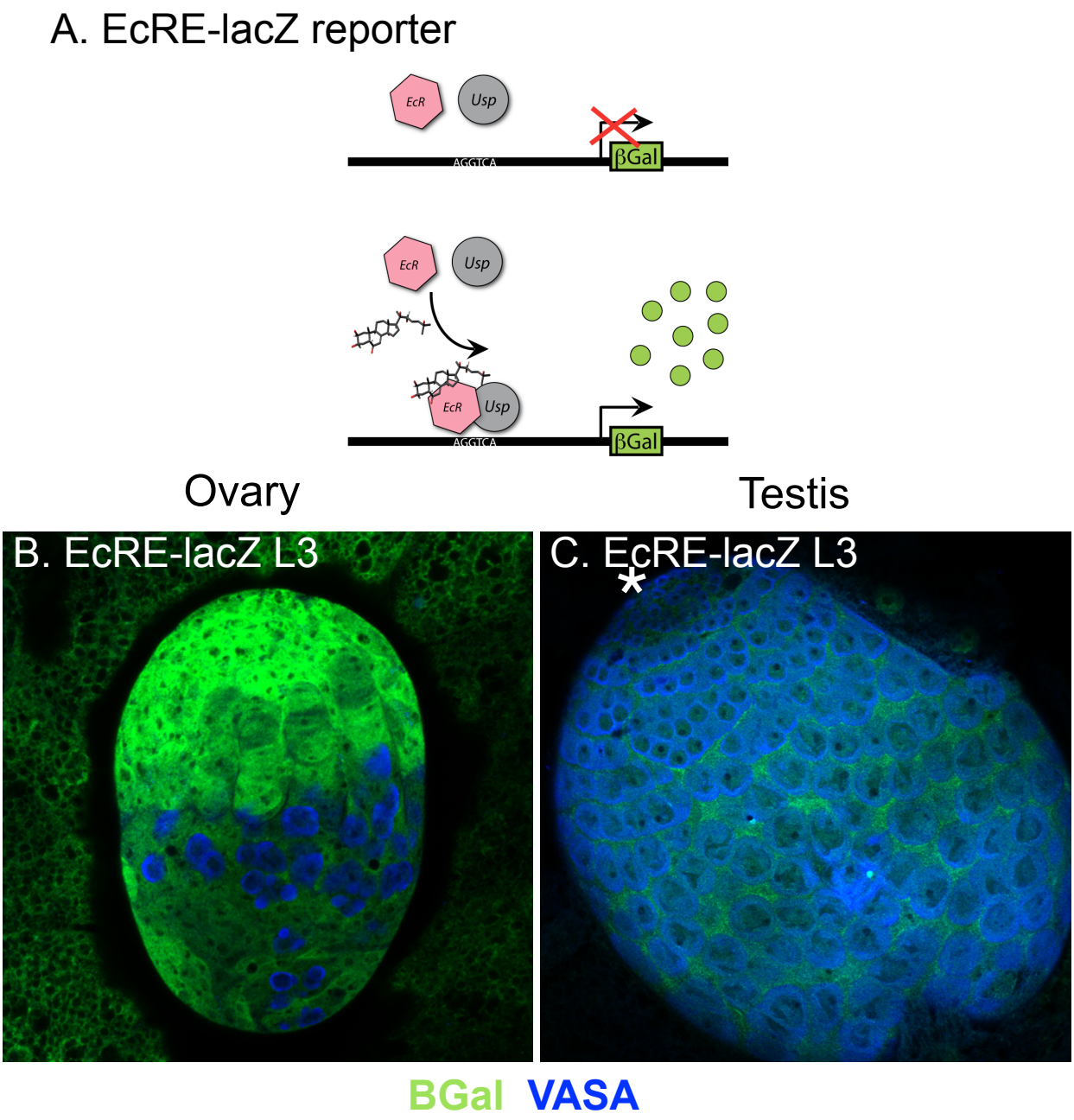


Figure 4.3. The larval ovary robustly responds to ecdysone signaling, but the testis does not. A) Schematic of EcRE-lacZ reporter system. Using a reporter for ecdysone receptor activity (ecdysone receptor DNA binding sites upstream of a lacZ reporter, EcRE-lacZ), we find that the larval ovary (B) shows robust activity while the larval testis shows little activity (C). Gonads dissected from larvae aged 120 hrs AEL.

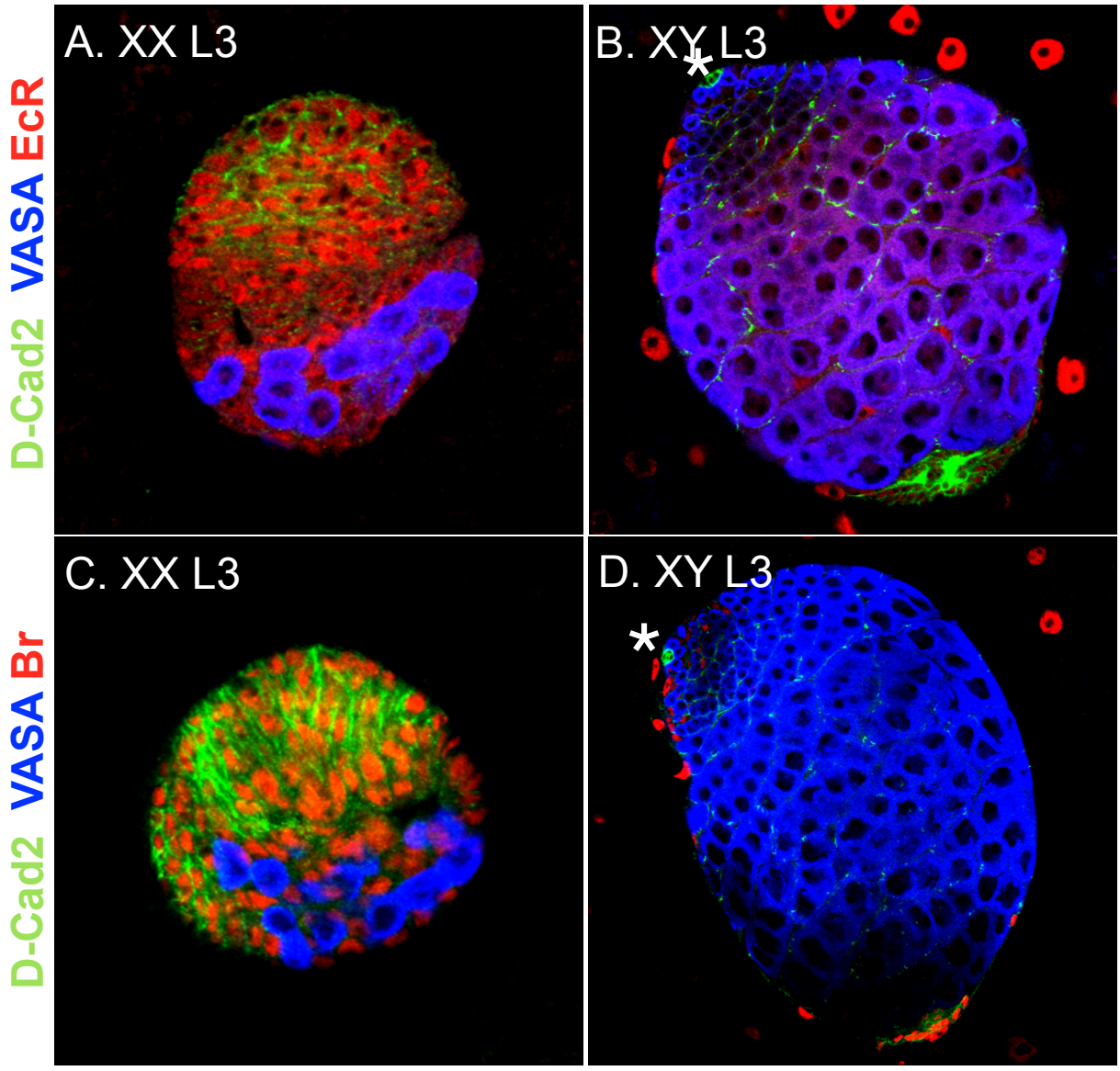


Figure 4.4. Ecdysone signaling is sexually dimorphic. A-B) Larval (L3) gonads, ovary (XX) and testis (XY), with germ cells marked with anti-Vasa (blue), anti-EcR (red), and D-Cad2 (green). C-D) Larval (L3) gonads, ovary (XX) and testis (XY), with germ cells marked with anti-Vasa (blue) and anti-Br (red), and D-Cad2 (green).

DSX regulation of EcR in the gonad

We predicted that the sex-differences in EcR expression and tissue responsiveness to ecdysone signaling could be regulated downstream of the sex-determination cascade, specifically by expression of sex-specific transcripts of *doublesex*, dsx^F and dsx^M in the gonad. In *dsx* mutants, some of both XX and XY animals end up with hubs, and others switch to form terminal filaments (Camara, Whitworth, and Van Doren, Unpublished). Thus, the XX animals do not robustly follow the female path, since many have hubs, and neither do the XY animals robustly maintain the male path, since many switch to form terminal filaments. We propose that the sex-differences we observe in ecdysone signaling is important for how *dsx* regulates the male and female paths in the gonads. If *dsx* modulates the ecdysone response so that it is robustly on in females, to promote ovary formation, and off in males, to allow for testis maintenance, then we expect in *dsx* mutants for there to be an intermediate ecdysone response in the gonads of both sexes (XX and XY).

We examined the effects of *dsx* mutations on the expression of EcR in gonads with a male niche component, the hub. In *dsx* mutant gonads with a male niche component, the hub, we detect EcR in all somatic nuclei of the gonad including the hub. Staining in germ cells was devoid of EcR (Figure 4.5A). This is in contrast to what we observed in wild-type larval testes with hubs that do not have EcR expression near the niche (Figure 4.4B). In *dsx* mutant gonads with a female niche component, the terminal filament, we also detect EcR staining in all the somatic nuclei of the gonad at comparable levels to wild-type larval ovaries. Again, EcR did not localize to germ cells in these gonads. Further, when we examined expression of the characteristic ecdysone responsive gene, *broad*, in *dsx* mutant gonads with a male niche component, the hub, we find an intermediate level of Broad immunostaining throughout the

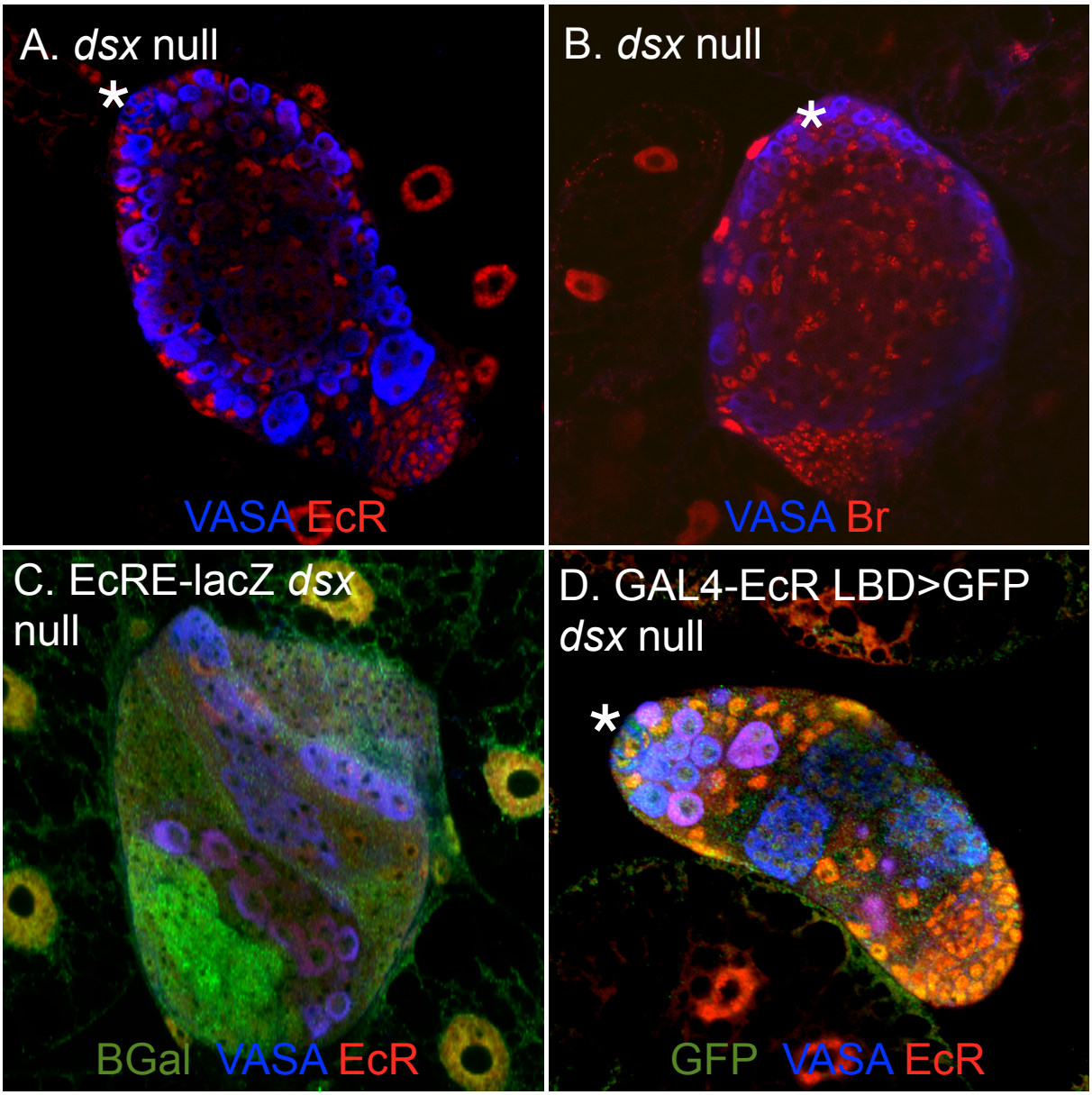
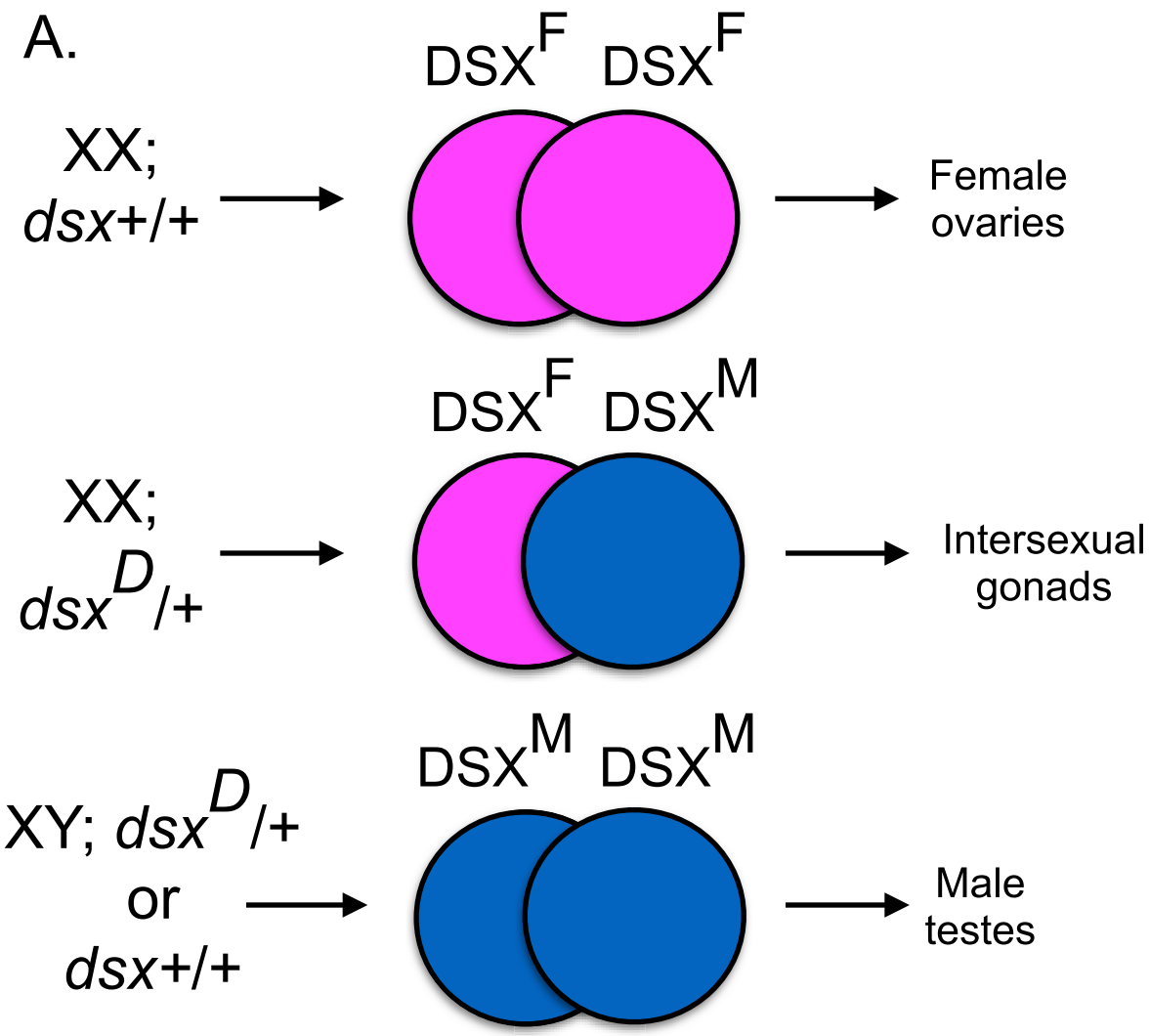


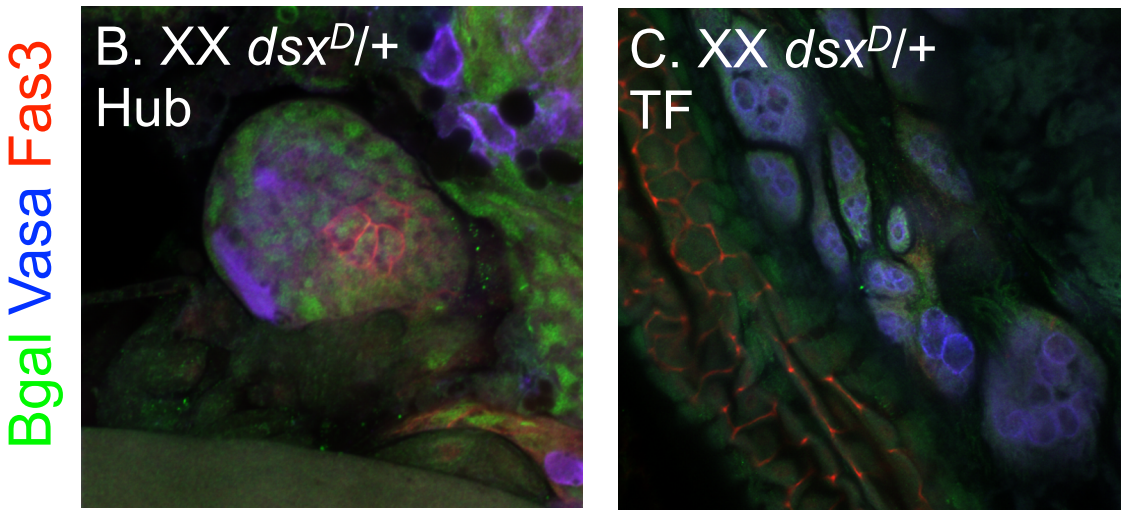
Figure 4.5. DSX regulates ecdysone signaling. A) L3 *dsx* null gonad with a hub. Germ cells marked with anti-Vasa (blue) and anti-EcR (red). B) L3 *dsx* null gonad with a hub. Germ cells marked with anti-Vasa (blue) and anti-Br (red). C) EcRE-lacZ *dsx* null larval testes with Bgal marked with anti-Bgal (green) and germ cells marked with anti-Vasa (blue). D) GAL4 EcR LBD>UAS-GFP.nls L3 *dsx* null gonad with germ cells marked with anti-Vasa (blue), GFP marked with anti-GFP (green), and anti-EcR (red).

gonad (Figure 4.5B). Thus, in contrast to sexually dimorphic expression observed in wild-type flies, lacking *dsx* function in gonads with hubs resulted in robust EcR expression. Since EcR expression was still observable in *dsx* mutant gonads with terminal filaments, these results revealed that the male specific DSX isoform, DSX^M, represses EcR. To further test this conclusion, we analyzed ecdysone receptor activity using a sensitized genetic background which produces only the male isoform of DSX regardless of chromosomal sex, dsx^D. In XX dsx^D/+ animals, the result of having both DSX^M and DSX^F in a single fly is thought to cause intersexual development similar to a *dsx* null (Figure 4.6A). Having active DSX^M and DSX^F protein in the same set of cells is thought to result in a DSX^M/DSX^F heterodimer and a cancelation of DSX function. We found that the level of EcR, which was elevated in all somatic cells of *dsx* mutant gonads with hubs, was also present in XX; dsx^D/+ gonads with hubs, confirming a repressive role of DSX^M on EcR expression (Figure 4.6BC). These data support the hypothesis that the steroid hormone ecdysone elicits a different response in the male vs. female gonad and that this difference is regulated by *dsx* and may be important for proper formation of the ovary vs. the testis.

To investigate the function of sex-specific ecdysone signaling, we performed genetic assays to test whether the ecdysone pathway exhibits the loss and gain of function phenotypes to support our model. We performed loss and gain of function experiments using a mutant background that results in intersexual development (dsx^D/+). As discussed earlier (Chapter 2), dsx^D/+ mutant animals express the male form of DSX, DSX^M, in the soma of both males and females. In XX dsx^D/+, gonads have an intersexual soma and are disorganized in morphology and develop either ovarioles with female-like niches (86% of gonads with terminal filaments), or a male-like niche (14% of gonads with hubs). To ensure that any phenotypes we observed are

Figure 4.6. DSX^M represses ecdysone signaling. A) XX; *dsx*+/+ animals only produce DSX^F (from *dsx*+). XX; dsx^D/+ animals produce both DSX^F (from *dsx*+) and DSX^M (from dsx^D) XY; dsx^D/+ animals only produce both DSX^M (from *dsx*+ & dsx^D). B) EcRE-lacZ reporter activity in adult XX; dsx^D/+ gonad with a hub. C) EcRE-lacZ reporter activity in adult XX; dsx^D/+ gonad with a terminal filaments. Bgal is marked with anti-Bgal (green), germ cells marked with anti-Vasa (blue), and hub marked with anti-fas3 (red).


A.
XX;
dsx+/+
DSXF DSXF
Female
ovaries
XX;
dsxD/+
DSXF DSXM
Intersexual
gonads
XY; dsxD/+
or
dsx+/+
DSXM DSXM
Male
testes


Bgal Vasa Fas3
B. XX dsxD/+
Hub
C. XX dsxD/+
TF

due to autonomous activity of *EcR* in the gonad, we knocked down or expressed *EcR* specifically in the gonad using gonad Gal4 drivers (C587-GAL4). We reasoned that if a EcR was required to promote female-like niche formation in the XX dsx^D/+ gonad, then knockdown of EcR in gonads from this fly would change the hub to terminal filament ratio and we would observe more male-like niches. In the opposite experiment, if EcR is required to promote female-like niche formation in the XX dsx^D/+ background, we reasoned that gain of function experiments of EcR in gonads from this fly would be more female-like. After modulating gene expression using loss and gain of function experiments, we examined the presence of a male-like niche component, hubs, or a female-like niche component, terminal filaments, based on morphological differences and markers in larval XX dsx^D/+ mutants (Figure 4.7).

Knockdown of EcR even under weak UAS-RNAi expression (18C) in the soma resulted in masculinization of the gonad (73% hubs, 19% terminal filaments, 8% no niche; n=26) in comparison to control animals expressing only the somatic gonad GAL4 driver (20% hubs, 74% terminal filaments, 6% no niche; n=35)(Figure 4.7). This masculinization confirms that wild-type *EcR* in important for female niche formation. In gain of function experiments where we over-expressed EcR in the somatic gonad of XX dsx^D/+ animals, we observed the same amount (75%) of terminal filaments being formed as controls. However, no gonads with hubs (0%) were observed and 25% of gonads were absent for a niche (n=20). This data suggests that although EcR over expression was insufficient to induce the formation of more terminal filaments, ectopic EcR expression is sufficient to repress the hub from forming in XX dsx^D/+ animals. Thus, *EcR* appears to be expressed in a sexually dimorphic manner during niche formation and is required to promote formation of the female niche over the male niche.

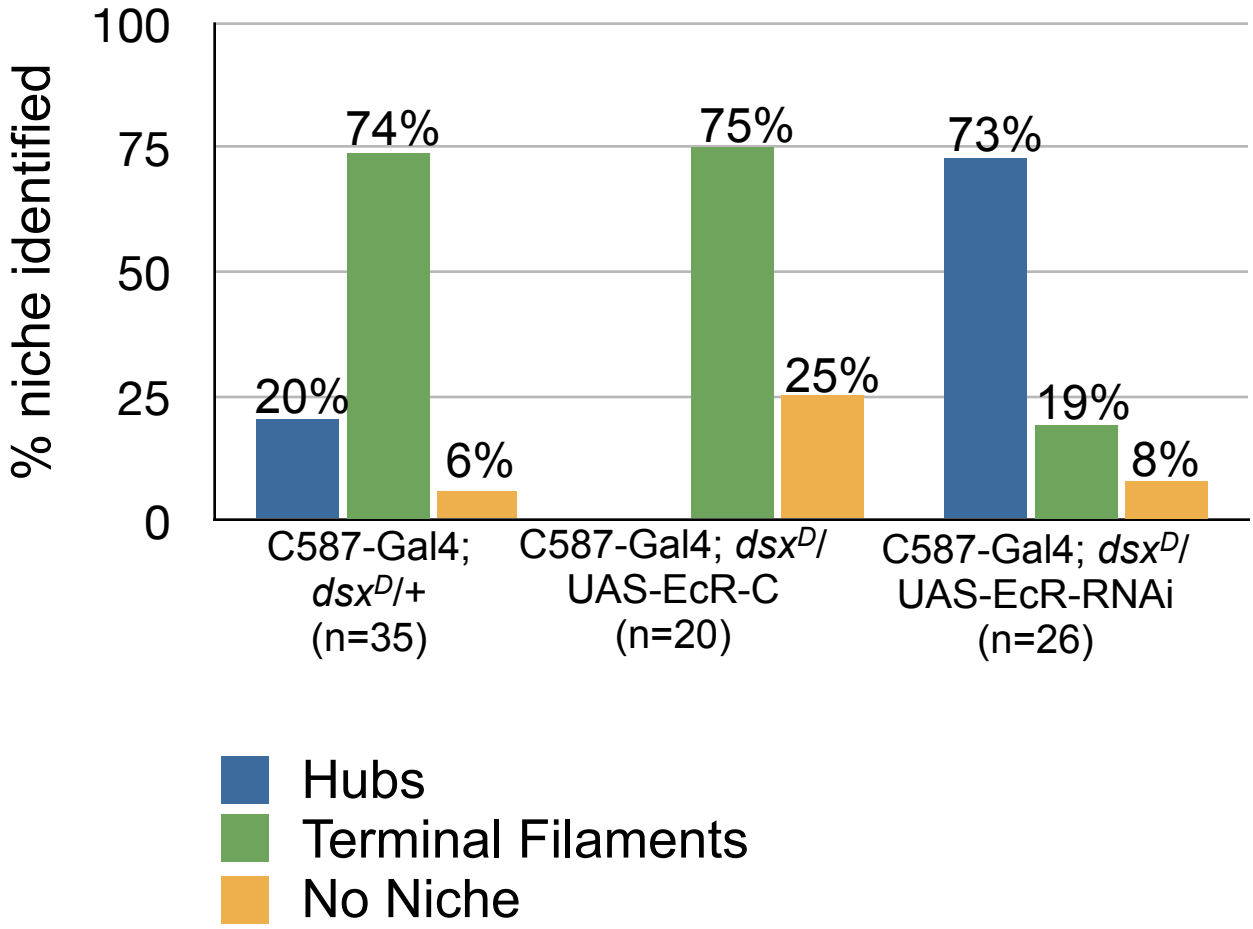


Figure 4.7. Function of dimorphic *EcR* in the gonad. Percentages of hubs (blue bar), terminal filaments (green bar), and absence of a niche (yellow bar) in gonads from animals expressing UAS-*EcR* or UAS-*EcR RNAi* in the somatic gonad of dsx^{D}/+ mutant animals.

Since our initial studies demonstrated that the EcR is sex-specifically expressed in wild-type larval ovaries vs. testes and that EcR promotes formation of the female niche over the male niche in XX dsx^{D}/+ mutant backgrounds, we sought to test for a functional role of EcR in causing differences in the gonads to respond to ecdysone signaling. To do this, we conducted EcR gain of function studies in a wild-type background. We predicted that since the ecdysone signaling pathway was sex-specifically higher in wild-type larval ovaries vs. testes and that EcR promoted formation of a female niche over the male niche in dsx mutant backgrounds, then driving expression of EcR in wild-type males might be sufficient to disrupt the male path and possibly to induce them to follow the female path.

We found that over-expression of EcR in the soma of the testis (*C587*-GAL4, Figure 4.10, or *Traffic jam*-GAL4, Figure 4.8), the hub (*Unpaired*-GAL4, Figure 4.11), and in *dsx* expressing cells (*dsx*-GAL4, Figure 4.9) is insufficient to activate ecdysone signaling in the testes. By examining a reporter for ecdysone receptor activity (EcRE-lacZ), we find that the larval testis shows little activity despite robust expression of EcR in the soma (Figure 4.9D). As expected only one region of the testis (the terminal epithelium), which is at the opposite end of the testis from the hub, shows EcRE-lacZ staining. Further, when we examined expression of *broad*, levels were not increased relative to control animals (Figure 4.8, 4.9, 4.10, 4.11).

Since the larval testis demonstrated to have lower ligand sensor activation in comparison to larval ovaries in earlier studies, one possibility is that the male gonad experiences lower ecdysone titers relative to the larval ovary. Mutants of ecdysone can be induced to pupariate by ecdysone supplementation in their food, thus we attempted to induce ecdysone signaling in testes by simultaneously expressing EcR in the soma of the testis using the UAS-GAL4 system and also growing the larvae on food supplemented with 1mg/mL 20-hydroxyecdysone (20E), an

active metabolite of ecdysone. By examining the reporter for ecdysone receptor activity (EcRE-lacZ), we find that the larval testis shows little activity even in the presence of supplemented 20E (Figure 4.12C). Further, when we examined expression of *broad*, protein levels were not increased relative to control animals (Figure 4.12D). Since neither over expression of EcR nor supplementation of the active metabolite of ecdysone can activate ecdysone signaling in the larval testis, we concluded that other components of the ecdysone signaling pathway are missing in the larval male testis (Figure 4.12). EcR can bind to the ecdysone hormone alone. However, in order for EcR to bind to ecdysone responsive elements and activate transcription of its target genes, EcR requires a co-factor known as Ultraspiracle (USP). In the presence of active ecdysone hormone, EcR dimerizes to USP and the EcR/USP/20E complex efficiently binds to ecdysone responsive elements to transcriptionally activates ecdysone responsive genes such as Broad. It is known that USP is expressed in the larval ovary and adult testis, but there is no literature indicating the presence of USP in the larval testis.

In addition to the gain of function studies, we have performed loss of function experiments and knocked down *EcR* activity specifically in the gonad using gonad GAL4 drivers (*Traffic jam*-GAL4) and UAS-*EcR RNAi* tested their affects in wild-type background. We confirmed that EcR activity is required during female development of ovaries and found that EcR is dispensable in male testes. The result that EcR is not required for normal development and function of the male gonad is consistent with previous research describing that EcR is dispensable for cyst cell development (Qian et al., 2014). In agreement with other experiments, larval ovaries expressing the dominant negative form of EcR isoforms (UAS-EcR-B1.W650A) in the somatic cells have fewer terminal filament cells are organized into stacks and primordial germ cells fail to intermingle with intermingled-cells (Gancz et al., 2011; Figure 4.13). On the

other hand, the testis exhibits remarkable normal morphology (Figure 4.13BD). Similar results were observed in knockdown experiments using UAS-RNAi and other dominant negative EcR alleles.

Overall, these studies allowed us to determine that the sex-specific ecdysone response that we have observed in the gonad has a functional consequence for sex-specific development of the gonad stem cell niche.

Figure 4.8. Driving expression of *EcR* in wild-type males in insufficient to disrupt the male path and induce them to follow a female path using the early somatic cell driver, *Traffic jam*-GAL4. A-A') Larval testes do not express EcR in the testis niche. B-B') Components of the male stem cell niche, the hub, persist when EcR is overexposed in the somatic cells of the testis (*Traffic jam-GAL4>UAS-EcR*). C-C') Broad, an early response gene of ecdysone signaling, is present at the onset of metamorphosis. D-D') Expression of EcR is insufficient to elevate levels of Br.

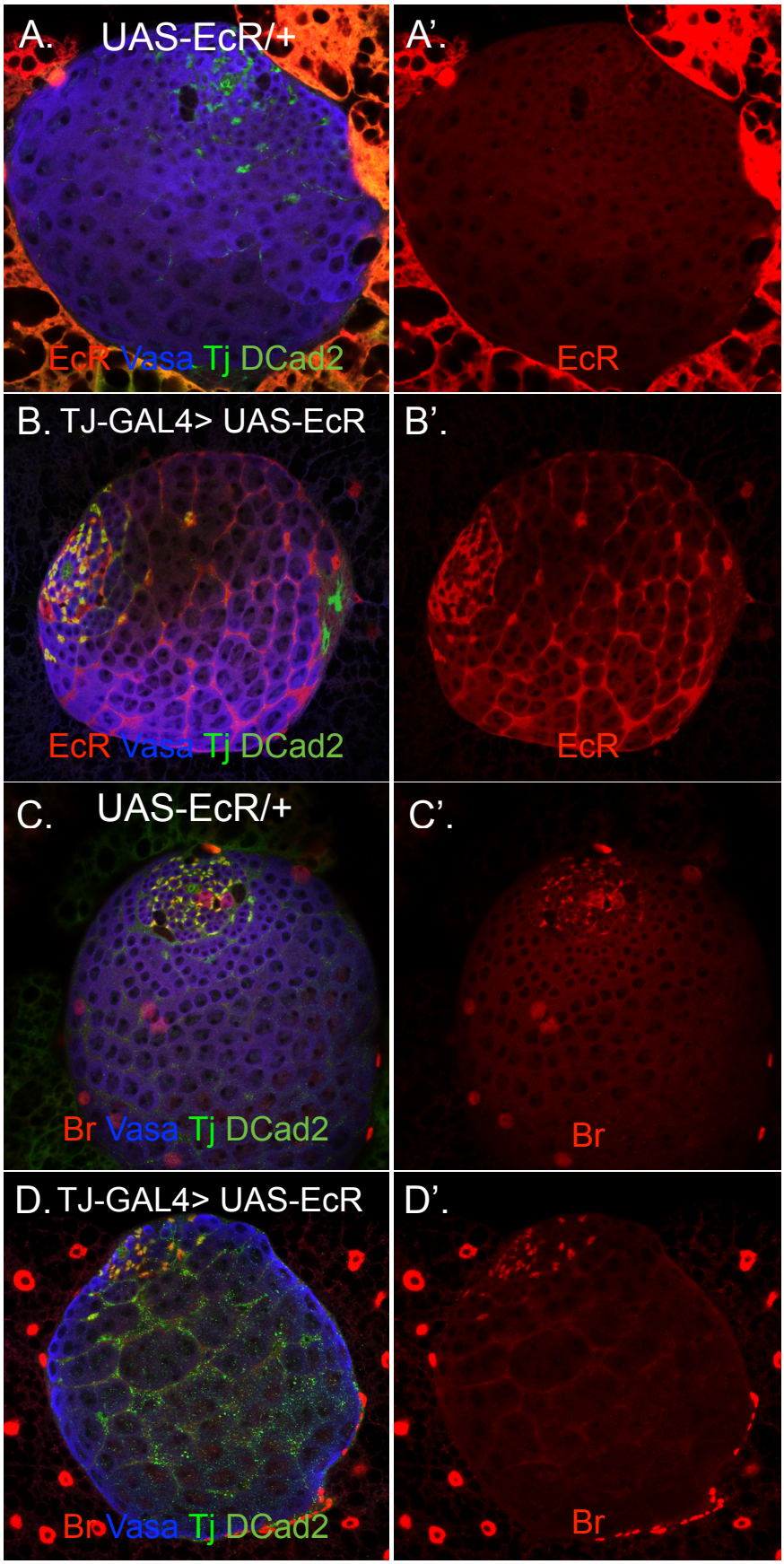

A. UAS-EcR/+
EcR Vasa Tj DCad2
A'.
EcR
B. TJ-GAL4> UAS-EcR
EcR Vasa Tj DCad2
B'.
EcR
C. UAS-EcR/+
Br Vasa Tj DCad2
C'.
Br
D. TJ-GAL4> UAS-EcR
Br Vasa Tj DCad2
D'.
Br

Figure 4.9. Driving expression of *EcR* in *dsx*-expressing cells wild-type males in insufficient to disrupt the male path and induce them to follow a female path. A-A') Components of the male stem cell niche, the hub, persist when *EcR* is overexpressed in the dsx expressing cells of the testis (*dsx*-GAL4>UAS-EcR). B-B'). Broad, an early response gene of ecdysone signaling, is present at the onset of metamorphosis in control testes (*dsx*-GAL4 only). C-C') Expression of *EcR* is insufficient to elevate levels of Br. D-D') Expression of *EcR* is sufficient to elevate EcRE-lacZ reporter activity in gonads expressing EcR (*dsx*-GAL4>UAS-*EcR*).

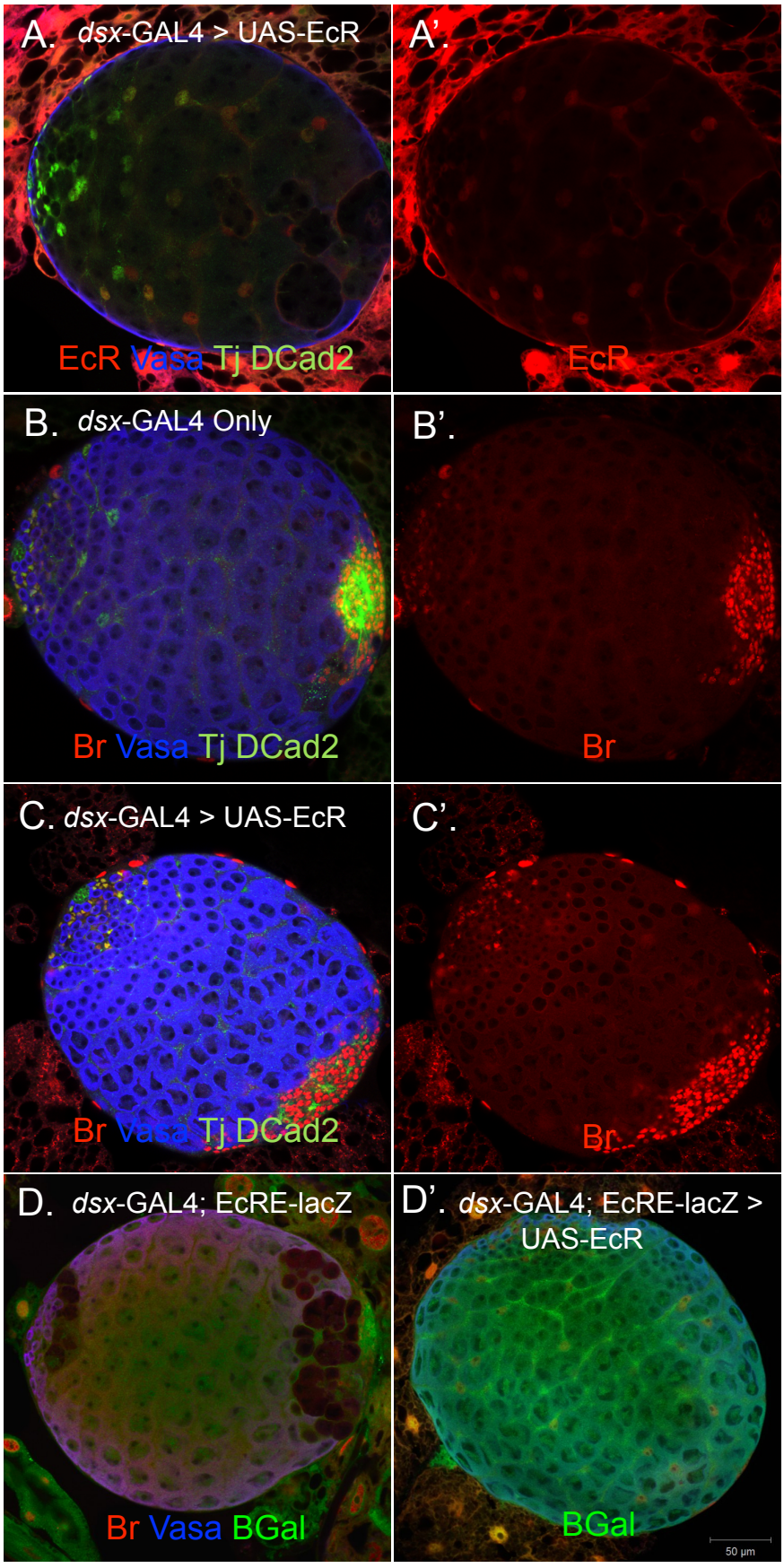
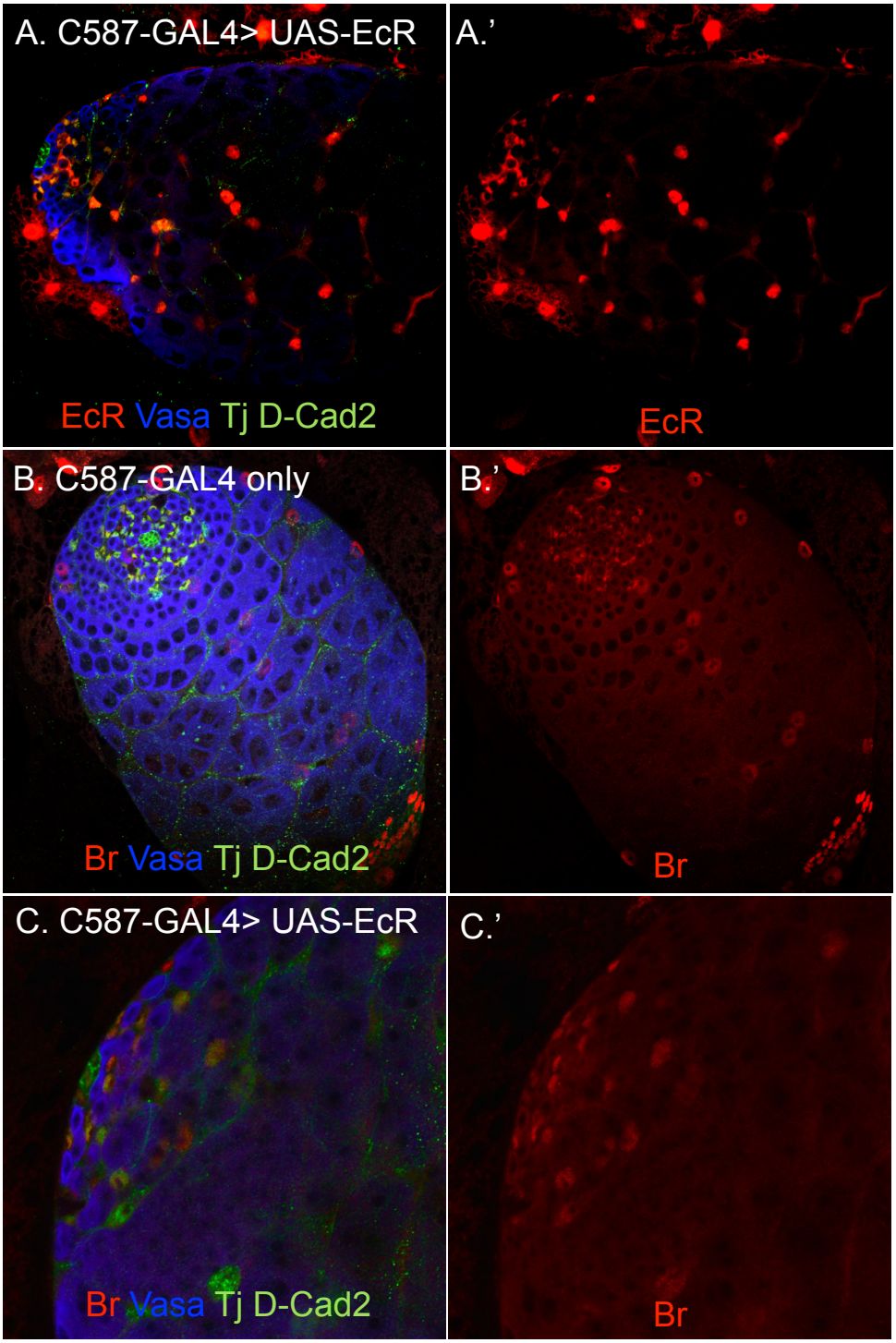

A. dsx-GAL4 > UAS-EcR
EcR Vasa Tj DCad2
A'.
EcR
B. dsx-GAL4 Only
Br Vasa Tj DCad2
B'.
Br
C. dsx-GAL4 > UAS-EcR
Br Vasa Tj DCad2
C'.
Br
D. dsx-GAL4; EcRE-lacZ
Br Vasa BGal
D'. dsx-GAL4; EcRE-lacZ > UAS-EcR
BGal
50 µm

Figure 4.10. Driving expression of *EcR* in wild-type males in insufficient to disrupt the male path and induce them to follow a female path using the somatic driver, *C587*-GAL4. A-A') Components of the male stem cell niche, the hub, persist when *EcR* is overexpressed in somatic cells of the testis (*C587*-GAL4>UAS-EcR). B-B'). Broad, an early response gene of ecdysone signaling, is present at the onset of metamorphosis in control testes (*C587*-GAL4 only). C-C') Expression of *EcR* is insufficient to elevate levels of Br.


A. C587-GAL4> UAS-EcR
EcR Vasa Tj D-Cad2
A.'
EcR
B. C587-GAL4 only
Br Vasa Tj D-Cad2
B.'
Br
C. C587-GAL4> UAS-EcR
Br Vasa Tj D-Cad2
C.'
Br

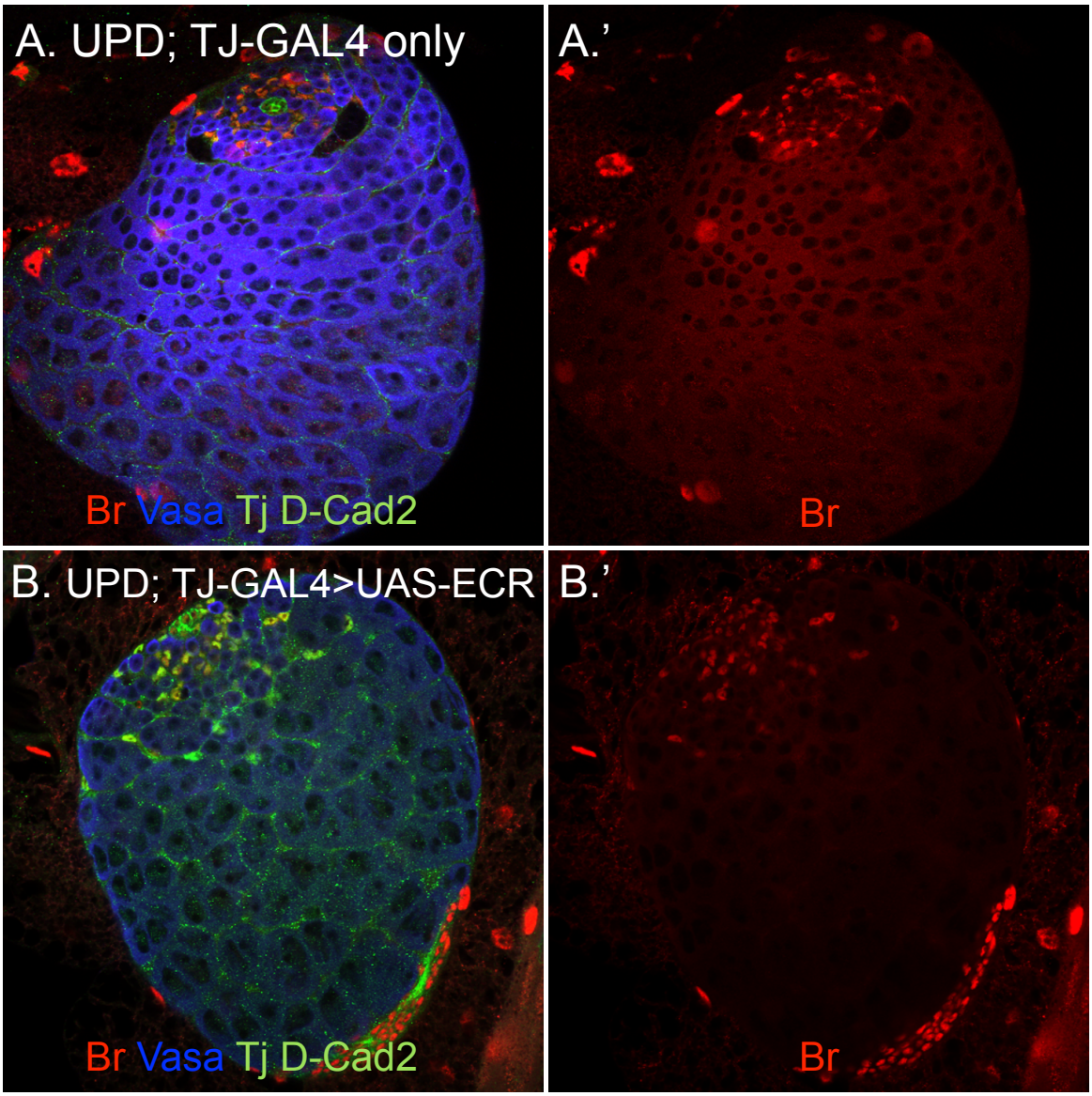


Figure 4.11. Driving expression of *EcR* in wild-type males in insufficient to disrupt the male path and induce them to follow a female path using the hub and early somatic cell drivers, *Unpaired*-GAL4 and *Traffic jam*-GAL4, respectively. Components of the male stem cell niche, the hub, persist when *EcR* is overexpressed in somatic cells of the testis (*Unpaired*-GAL4 (UPD); *Traffic jam*-GAL4>UAS-EcR). A-A'). Broad is present at the onset of metamorphosis in control testes (*Unpaired*-GAL4 (UPD); *Traffic jam*-GAL4 only). C-C') Expression of *EcR* is insufficient to elevate levels of Br (*Unpaired*-GAL4 (UPD); *Traffic jam*-GAL4>UAS-*EcR*).

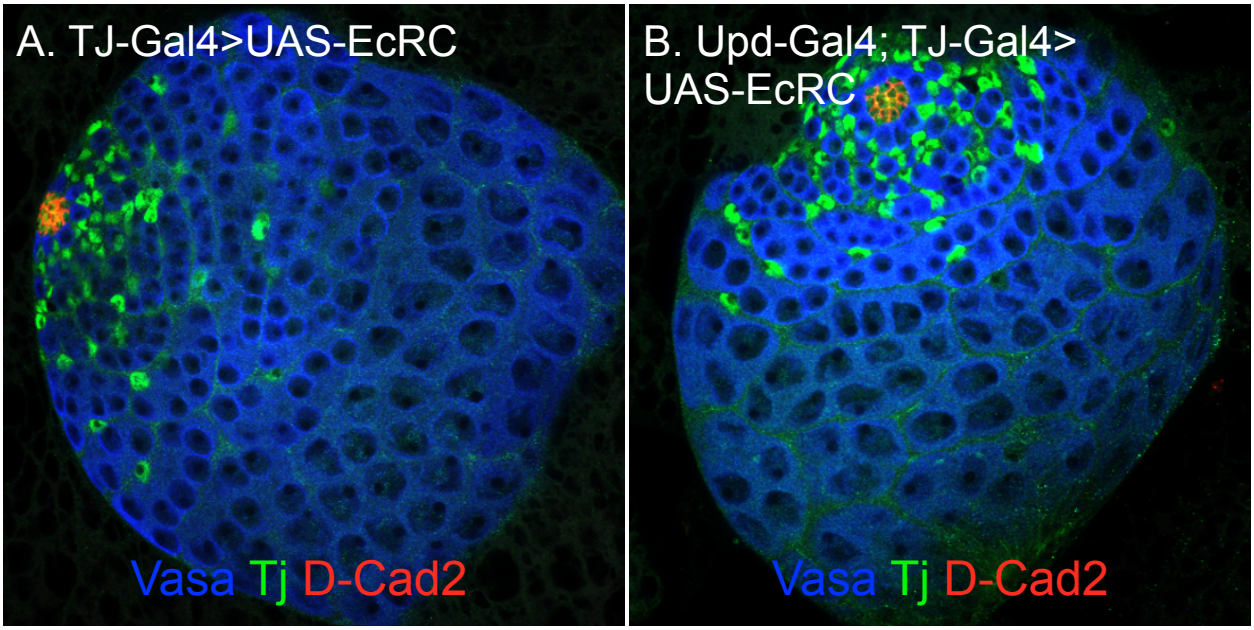


Figure 4.12. 20-Hydroxyecdysone feeding and driving expression of *EcR* in wild-type males in insufficient to disrupt the male path and induce them to follow a female path. Components of the male stem cell niche, the hub, persist when EcR is overexpressed in somatic cells of the testis (*Unpaired*-GAL4 (UPD); *Traffic jam*-GAL4>UAS-*EcR*) and in the presence of supplemented 20-hydroxyecdysone. A). Testes from larval males expressing *EcR* (*Unpaired*-GAL4 (UPD); *Traffic jam*-GAL4 only). B) Testes from larval males expressing *EcR* (*Traffic jam*-GAL4>UAS-*EcR*).

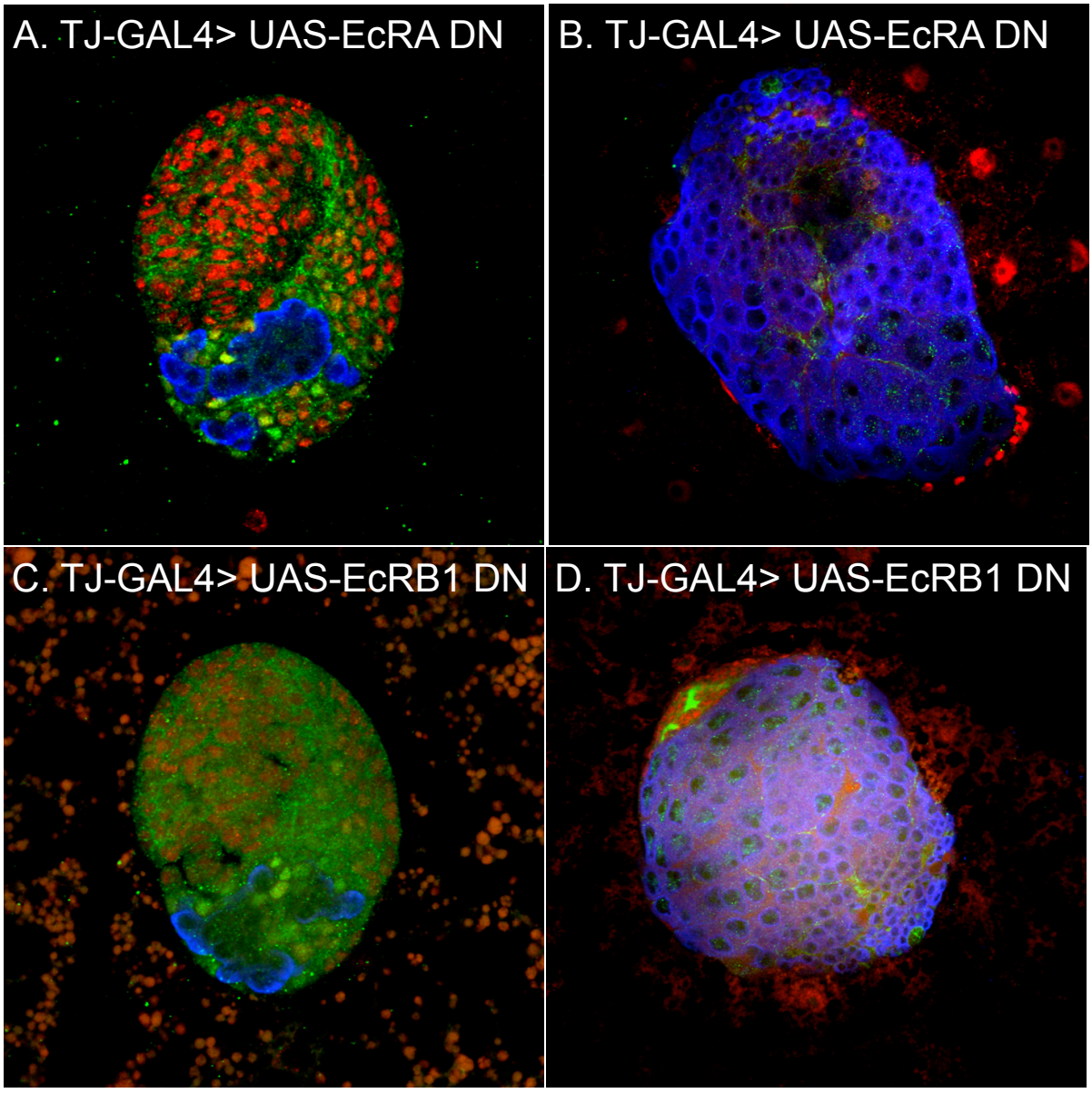


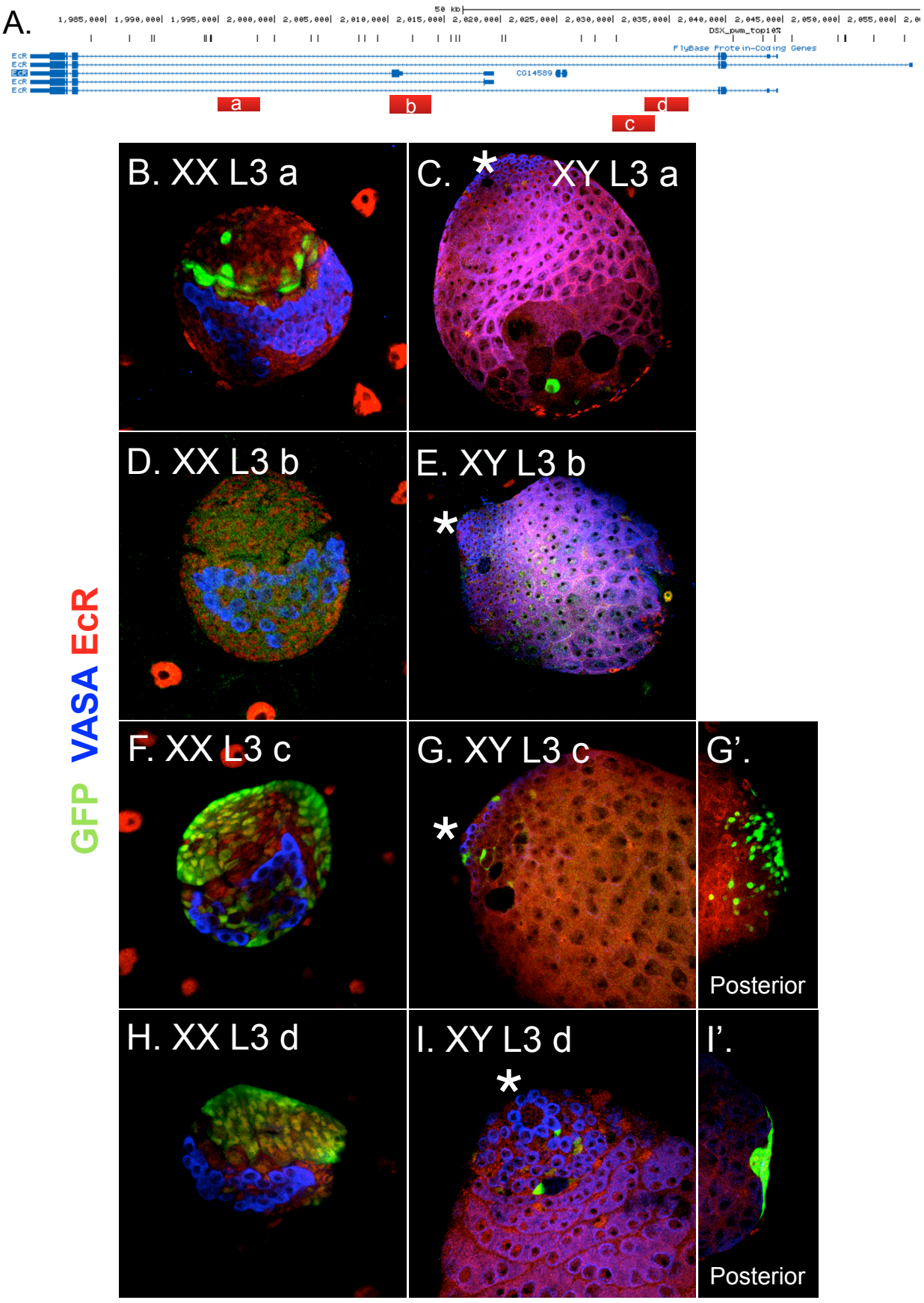
Figure 4.13. Knockdown of *EcR* results in a female specific phenotype. Expression of *EcR* dominant negative alleles (DN) can compete with endogenous EcRs resulting in a knockdown of ecdysone-induced gene activation upon expression. Expression of either dominant negative form of EcRA or EcRB1 (*TJ*-GAL4> UAS-*EcRA DN* or *TJ*-GAL4>UAS-*EcRB1* DN) has no effect on the development of the male testis, but results in defective ovarian morphogenesis.

DSX is a direct regulator of the ecdysone response

As indicated above, our genomic analysis has identified *EcR* as a direct DSX target. We have whole-genome DSX binding data for different tissues, as well as an evolutionary analysis of DSX consensus binding sites in different *Drosophila* tissues (Chapter 2). There are several regions of the *EcR* upstream sequence that show occupancy for the DSX protein, and one promoter in particular (P2) that has six positions of high scoring (>90th percentile PWM score) DSX binding sites that are conserved in at least one other *Drosophila* species (Figure 4.15A). This region is an outstanding candidate for direct regulation by DSX.

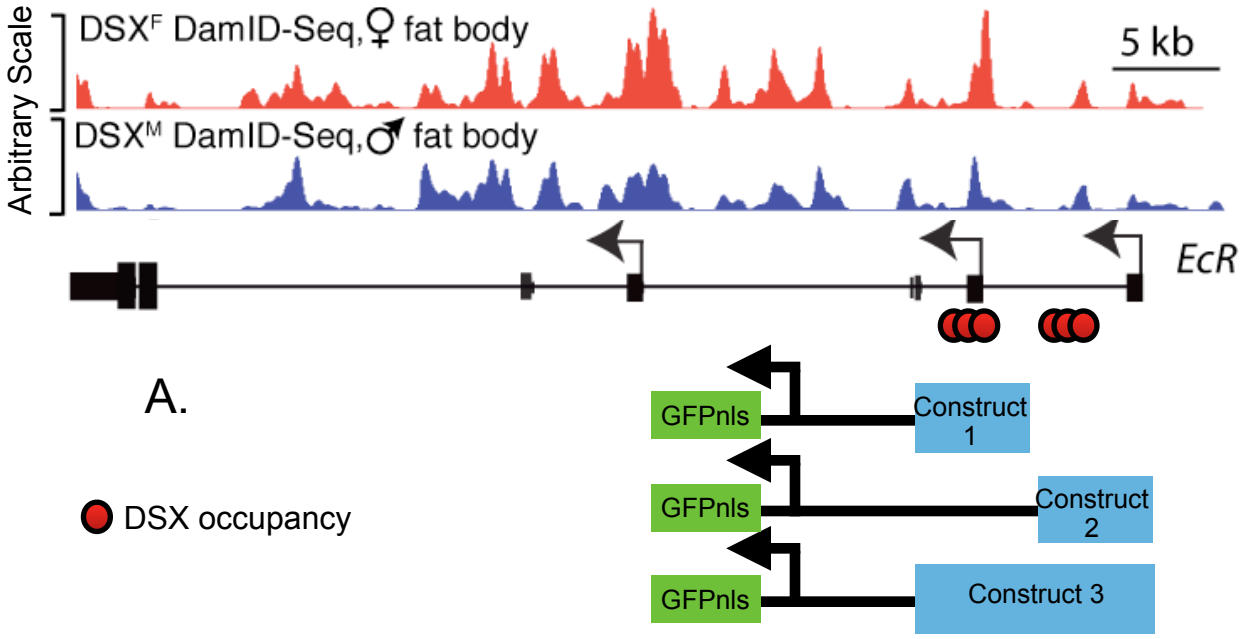
In addition to generating larger enhancer constructs to recapitulate *dsx*-dependent gonad expression, we tested a few smaller enhancer constructs available from the community (Janelia Farms Brain Enhancers). Janelia Farm GAL4 lines express GAL4 under control of defined sequence fragments from intronic regions of EcR (Figure 4.14A; red boxes; not drawn to scale). Although none of the Janelia Farm EcR GAL4 lines do not overlap with any of the high scoring DSX binding sites, we examined larval ovaries and testes for the presence of enhancers of EcR (Figure 4.12): Lines 'a' (line 46268), 'b' (line 45719), 'c' (line 48167), and 'd' (50200). We observed GFP expression in larval ovaries and testes from all four of the tested EcR enhancer GAL4 lines, but none recapitulated what we observe by EcR antibody staining. In the larval ovary, three of the four lines reported expression that did not recapitulate EcR expression detected by antibody staining (Figure 4.14BFH). In the larval testis, two of the four lines reported expression in the early somatic cells and terminal epithelium of the testis which is in contrast to what is observed by EcR antibody staining (Figure 4.14GI). Since none of the Janelia Farm EcR Gal4 lines recapitulated what is observed by EcR antibody staining, we did not follow up on any of the lines to test for *dsx*-dependence.

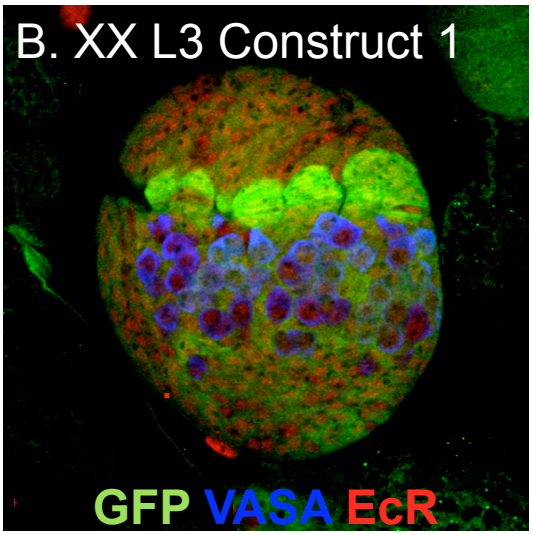
Figure 4.14. Enhancer constructs recapitulate sex-specific EcR expression in larval gonads. A) UCSC Browser Shot of *EcR* loci with positions of high scoring DSX binding sites, within 1kb upstream of genes or within the gene body and within the top percentile of PWM scores (~17k sites). (Clough, E., Jimenez, E., Whitworth, C., Kim, Y., et al, 2014). Janelia Farm GAL4 lines express GAL4 under control of defined sequence fragments from intronic regions of EcR (red boxes; not drawn to scale). Janelia Farm EcR GAL4 lines do not overlap with any of the high scoring DSX binding sites. Janelia Farm GAL4 lines 'a' (line 46268), 'b' (line 45719), 'c' (line 48167), and 'd' (50200) were crossed to UAS-*GFPnls*. B) Line 'a' reports expression in ovaries and testes. In the larval ovary, Line 'a' is specifically expressed in every terminal filament cell and a subset of somatic cells in the anterior soma. C) In the larval testis, line 'a' is lowly expressed in the posterior of the testis and in fat body. D) Line 'b' expression is not present in the larval ovary. E) In the larval testis, line 'b' is expressed in spermatocytes and pigment cells. F) In the larval ovary, line 'c' is expressed throughout the ovary in the anterior soma, a subset of terminal filament cells, intermingled cells and basal cells. This line recapitulates EcR expression detected by antibody staining. G) Line 'c', is expressed in the early somatic cells and in the posterior of the testis (G'.). H) In the larval ovary, line 'd' is lowly expressed in terminal filaments but highly expressed in the anterior soma. I) In the testis, line 'd' is in a few early somatic cells and is highly expressed in the somatic cells of the posterior of the testis (I'.)

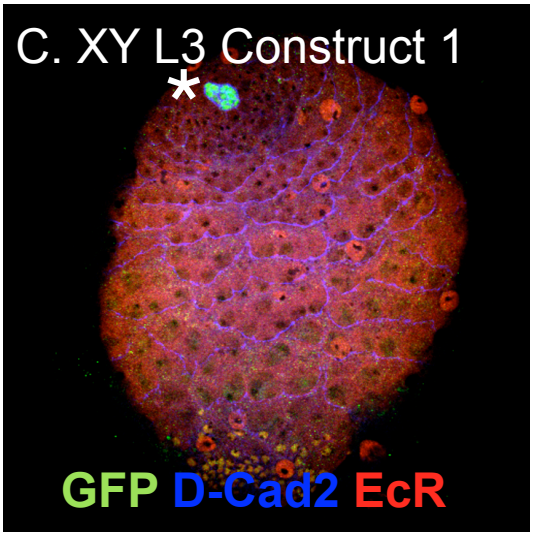

A.
50 kb
1,985,000 1,990,000 1,995,000 2,000,000 2,005,000 2,010,000 2,015,000 2,020,000 2,025,000 2,030,000 2,035,000 2,040,000 2,045,000 2,050,000 2,055,000
DSX_pwm_top10%
FlyBase Protein-Coding Genes
EcR
EcR
EcR
EcR
EcR
CG14589
a
b
c
d
B. XX L3 a
C. XY L3 a
D. XX L3 b
E. XY L3 b
F. XX L3 c
G. XY L3 c
G'.
Posterior
H. XX L3 d
I. XY L3 d
I'.
Posterior
GFP VASA EcR

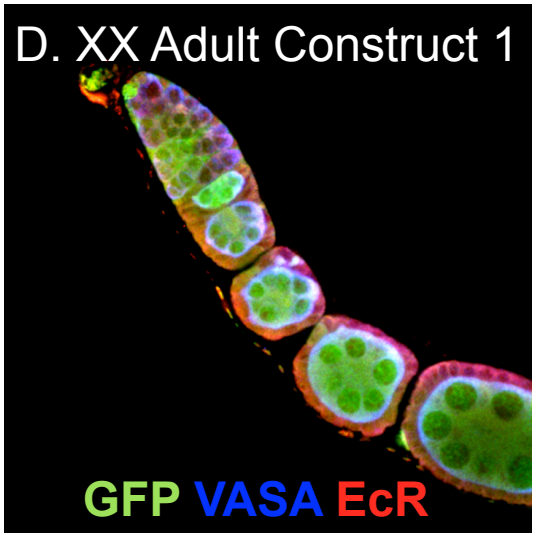
We used this data to identify putative sex-specific gonad enhancers for *EcR*, and generated two enhancer-reporter constructs, 'Construct 1 and 2', for these regions (using GFP reporter vectors and site-directed integration into the Drosophila genome). 'Construct 1' contains a perfect match to the DSX consensus site and three conserved DSX binding sites, while 'Construct 2' contains three positions of high scoring DSX binding sites that are conserved (see cartoon of constructs in Figure 4.15A). We examined reporter expression in both wild-type and mutant background that results in intersexual development (dsx^D/+) (Figure 4.15, Figure 4.16). When we examined Construct 1 in wild-type larval gonads, we find that the construct is expressed specifically in hubs from male larval testes and in terminal filaments of the larval testes (Figure 4.15BC, Figure 4.16AB). In addition to the terminal filament cells of the larval ovary, Construct 1 is expressed in the anterior soma, intermingled cells, and basal cells but at lower levels (Figure 4.13B). In adult female ovaries, Construct 1 expression persists in the terminal filament cells and expression is observed in the cap cells, escort cells, and in germ cells (Figure 4.15D). Construct 1 appears to recapitulate EcR staining for larval and adult ovaries. In the larval testis, Construct 1 is robustly expressed in the hub cells and weakly expressed in the posterior of the testis, the terminal epithelium (Figure 4.15C, Figure 4.16B). In the larval and adult, the presence of Construct 1 in the hub does not recapitulate EcR staining (Figure 4.15CE). Thus, Construct 1 exhibits tissue-specific activity, but lacks sex-specific activity. When we examined this construct in XX dsx^D/+ mutant gonads with hubs or terminal filaments to determine whether expression of Construct 1 is *dsx*-dependent, the expression of this reporter is abolished (Figure 4.15CD). This result suggests that DSX^M plays an active role in repressing expression of Construct 1. Thus, tissue-specific expression of Construct 1 requires *doublesex* function. Future experiments involve examining Construct 2 and 3 and following up on

Figure 4.15. *EcR* is occupied by DSX. A) Chromosomal map from genomic regions of EcR loci bound by DSX identified by DamID-seq on male and female fat body tissues (Clough, E., Jimenez, E., Whitworth, C., Kim, Y., et al, 2014). Using this data, we identified putative sex-specific gonad enhancers for *EcR* and generated three enhancer-reporter constructs for these regions using GFP reporter vectors and site-directed integration into the Drosophila genome. Only Construct 1 was evaluated in this work. B-E) Construct 1 GFP expression in a wild-type background. B) During larval development of the ovary, Construct 1 reports robust GFP expression in terminal filaments. Lower levels of expression are found in the anterior soma, intermingled cells, and basel cells. No expression is found in germ cells. C) In the larval testis, Construct 1 reports cell specific expression in every cell of the hub. The only other region with GFP is the terminal epithelium, the posterior of the testis. D) In adult ovaries, Construct 1 GFP expression is found in terminal filaments, cap cells, escort cells, and in germ cells. E) In adult testes, Construct 1 expression persists in the hub cells.


Arbitrary Scale
DSX^F DamID-Seq, ♀ fat body
5 kb
DSX^M DamID-Seq, ♂ fat body
EcR
A.
DSX occupancy
GFPnls
Construct 1
GFPnls
Construct 2
GFPnls
Construct 3


B. XX L3 Construct 1
GFP VASA EcR


C. XY L3 Construct 1
*
GFP D-Cad2 EcR


D. XX Adult Construct 1
GFP VASA EcR

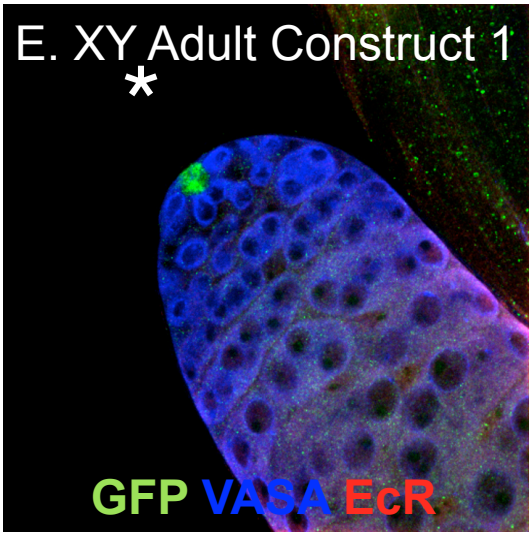
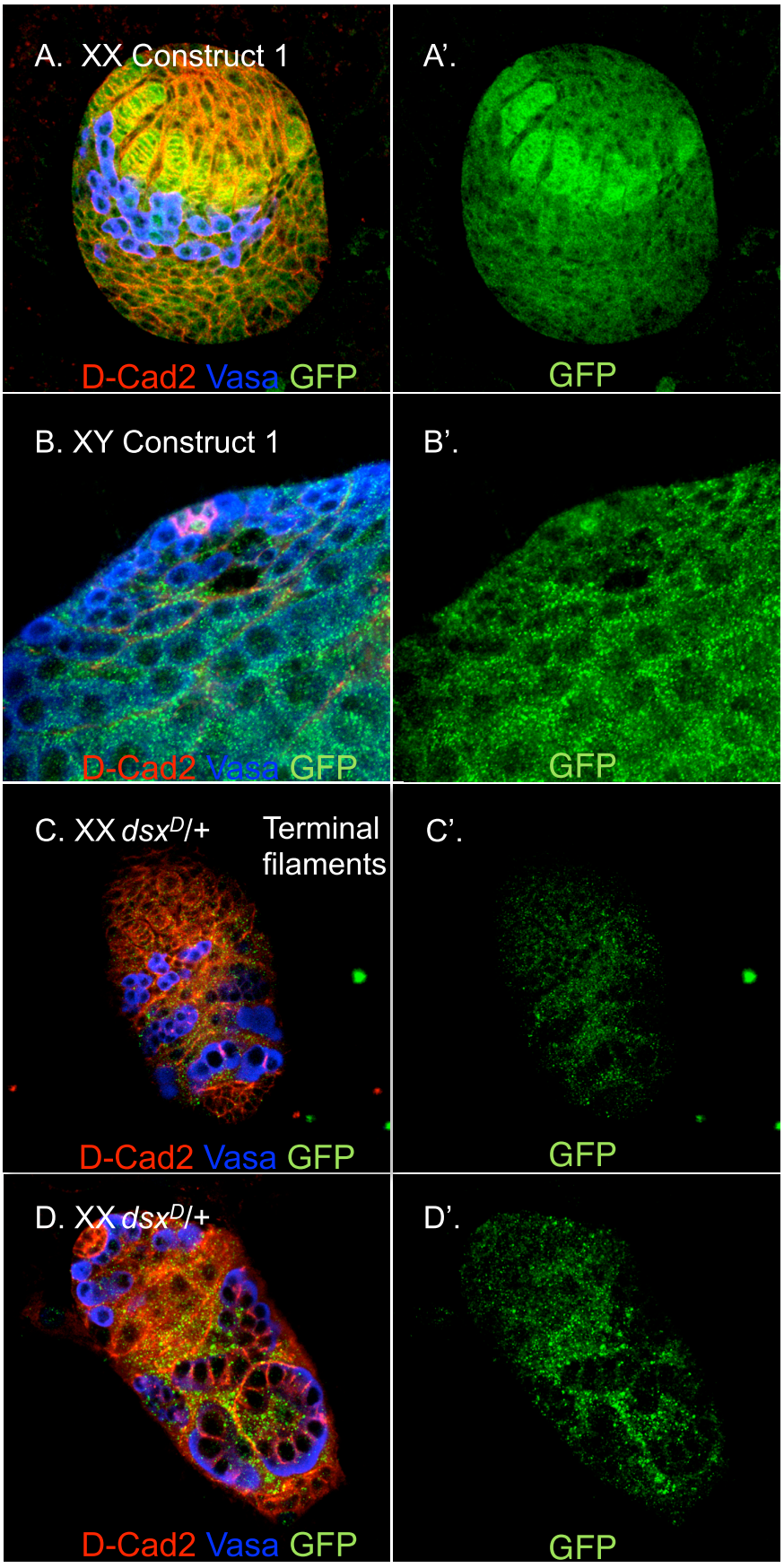

E. XY Adult Construct 1
*
GFP VASA EcR

Figure 4.16. *EcR* elements are regulated by *dsx in vivo*. Construct 1 (EcR-3.5k-GFP) reporter expression in the XX larval ovary (A) and XY larval testis (B). Shown is a channel of GFP only for both the larval ovary (A') and testis (B'). Construct 1 (EcR-3.5k-GFP) reporter expression in XX dsx^{D}/+ gonads with terminal filaments (C) and with a hub (D). Shown are GFP only channels as indicated.


A. XX Construct 1
D-Cad2 Vasa GFP
A'.
GFP
B. XY Construct 1
D-Cad2 Vasa GFP
B'.
GFP
C. XX dsx^D/+
Terminal filaments
D-Cad2 Vasa GFP
C'.
GFP
D. XX dsx^D/+
D-Cad2 Vasa GFP
D'.
GFP

Construct 1 in a *dsx* null background to determine whether it is truly *dsx*-dependent. Since my previous data indicates that EcR expression is high in females, extremely low in males and persists in *dsx* mutants, then we would expect the same behavior from our enhancer construct. Since our initial construct exhibits tissue-specific and dsx^M dependent activity, we have generated a new enhancer construct where we have mutated the consensus DSX binding sites in this construct (Figure 15A). This new construct has already been injected for site-directed integration into the *Drosophila* genome and is being screened for transformants carrying the construct. Through this analysis, we will be able to determine if the *EcR* gene is a direct target for regulation by DSX. Few DSX target genes are known, and so this analysis will be significant both from the perspective of sex-specific steroid signaling and in terms of how DSX acts as a transcriptional regulator of sex-specific gene expression.

Discussion

dsx acts to "tip the balance" between male and female developmental programs

An important question about the creation of sexual dimorphism is how a key transcription factor like DSX regulates a sex-specific developmental program. The formation of sex-specific structures, such as the hub or terminal filaments, can still form in the complete absence of *dsx* independently of the chromosomal constitution of the animal (XX vs. XY). The hubs and terminal filaments that form in *dsx* mutants have many of the characteristics of the wild-type structures, including the proper morphology and pattern of gene expression, and they associate with and signal to the germ cells normally (Camara, unpublished). DSX does not independently activate the many different genes that are likely to be required to form these structures. Instead, DSX acts to ensure that the male structures form reliably in XY animals, while the female

structures form in XX animals. To achieve this, DSX is likely to regulate male- or female-specific developmental programs that can then function independently of DSX. Thus, DSX would act to "tip the balance" between whether the male (hub) or female (terminal filaments) pathway was activated. One way in which DSX might regulate this balance is by influencing the expression of key upstream regulators of these pathways, such as ecdysone signaling. From this work, ecdysone signaling represents the best candidate for a "pro female gene" in our model of sex-specific gonad niche development (Figure 4.17).

Ecdysone signaling response in female vs male gonads

Ecdysone signaling activity in the larval female gonad, but not the male indicates that the gonads have dimorphic components involved in activation of ecdysone signaling. Differences in how, where, and when sex-specific gonads respond to ecdysone signaling may be attributed to male and female gonads carrying out different functions. Even though other tissues in the male must respond to ecdysone to control metamorphosis, the testis remains much as it has since the end of embryogenesis with continued spermatogenesis.

Ecdysteroid levels in adult males have been shown to be lower compared to the high-level pulses produced by the prothoracic gland that drive transitions during development and those of adult female insects. In male adults, genes important for the biosynthesis of ecdysone are lower in expression compared to larvae, pupae, and females (Hentze et al., 2013). Since identification of ecdysteroidogenic tissue relies on the ability of tissue to convert labeled C to E and 20E, the low ecdysteroidogenic capacity of males makes it difficult to detect conversions in the adult (Hentze et al., 2013). Indeed, when we examined the spatial patterns of ecdysone ligand sensor activation in male and female gonads at the onset of metamorphosis, we detected activation in the somatic cells of the ovary at higher levels compared to the testis. Thus, we show

here that male and female gonads respond differently to the late larval ecdysteroid pulse. One explanation for this is that in addition to having lower ecdysteroid levels in larval male testes compared to larval ovaries, the larval testis also has lower levels of expression of the EcR co-activator protein Ultraspiracle (USP). EcR can bind to the ecdysone hormone alone. However, in order for EcR to bind to ecdysone responsive elements and activate transcription of its targets, EcR requires a co-factor, USP. In the presence of ecdysone hormone, EcR dimerizes to USP and the EcR/USP/20E complex binds to ecdysone responsive elements to transcriptionally activate genes. Colocalization studies have demonstrated that USP colocalizes with EcR in all tissues and stages that have been examined thus far (Hodin and Riddiford, 1998; Yao et al., 1993). It is well known that USP is expressed in the ovary and in some cells of the adult testis, but there is a lack of data indicating the presence of USP in the larval testis (Yi., et al., 2014). Future experiments involve determining the expression status of USP using a green fluorescent protein tagged protein at the native genomic USP locus followed by gain of function studies by expressing the EcR/USP complex to efficiently activate the ecdysone signaling pathway in the testis. These studies will further determine if the sex-specific ecdysone response that we have observed in the gonad has a functional consequence for sex-specific development of the gonad.

In support of this, we have found that males and females exhibit a dramatically different response to ecdysone in the gonads. By examining a reporter for ecdysone receptor activity (ecdysone receptor DNA binding sites upstream of a lacZ reporter, EcRE-lacZ), we find that the larval ovary shows robust activity of this reporter while the same stage larval testis shows little activity. Only one region of the testis (the terminal epithelium), which is at the opposite end of the testis from the hub, shows EcRE-lacZ staining. This internal control allows us to conclude

that the testis is able to respond to ecdysone signaling, but most of the testis, including the gonad stem cell niche, just does not.

The most straightforward way to have female-specific ecdysone response would be to have the ecdysone receptor (*EcR*) expressed only in the female and not in the male gonad. Indeed, we find EcR in all somatic nuclei of the larval ovary, but not the testis. We hypothesize that the sex-differences we observe in ecdysone signaling is important for how *dsx* regulates the male and female paths in the gonads. If *dsx* modulates the ecdysone response so that it is robustly on in females, to promote ovary formation, and off in males, to allow for testis maintenance, then we expect in *dsx* mutants for there to be an intermediate ecdysone response in the gonads of both sexes (XX and XY). In support of this model we detect EcR in all somatic nuclei of *dsx* mutant gonads with a male niche component, the hub. These data suggest that the male form of *dsx*, dsx^M, is important for repressing EcR expression and ecdysone signaling in the testis. Whereas, the female form of *dsx*, dsx^F, is important for ensuring female structures form in XX animals. Thus, the steroid hormone ecdysone elicits a different response in the male vs. female gonad and this difference is regulated by *dsx* which is important for proper formation of the ovary vs. the testis.

We show that knockdown of EcR not only inhibits terminal filament formation, but can promote hub formation. We also show that over-expression of EcR, can repress the hub from forming. It appears that *Ecdysone receptor* acts as a "pro female gene" and that *dsx* allows the steroid hormone ecdysone to act differently in males and females, so that males ignore this signal and maintain the hub and testis, while females produce a robust response to form ovaries.

Our genomic analysis indicates that *EcR* exhibits the hallmarks of being a DSX target (Chapter 1 or 2). We have whole-genome DSX binding data for different tissues, as well as an

evolutionary analysis of DSX consensus binding sites in different Drosophila tissues. There are several regions of the *EcR* upstream sequence that show occupancy for the DSX protein, and one promoter in particular (P2) that has six positions of high scoring (>_ 90th percentile PWM score) DSX binding sites that are conserved in at least one other *Drosophila* species. Here we show that there is at least one region of the EcR locus that is expressed in a tissue-specific manner and that this expression in the gonad is regulated by DSX.

Parallels between flies and mammals

An established concept in developmental biology concerns the differences in sexual differentiation of the gonads in mammals compared to insects. In mammals, sexual differentiation of the gonad depends on gonad specific expression of the SRY gene on the Y chromosome. In the presence of SRY expression, the bipotential gonad will develop into a testis and in its absence develops into an ovary. The end result are sex specific gonads that secrete sex hormones which direct development of the sexual characteristics that differentiate the sexes. Like mammals, sex determination in Drosophila is initialized by the presence of two X chromosomes in females and XY in males. Rather than utilizing sex hormones to direct the characteristics that distinguish males from females, sexual dimorphism in the fly is regulated by the doublesex gene. The role of *doublesex* is highly conserved in different insects and *dsx* homologs (*dsx*, *mab-3* related transcription factors, DMRTs) play roles in sexual differentiation in a diverse array of metazoans.

Our data supports the hypothesis that the steroid hormone ecdysone elicits a different response in the male vs. female gonad and that this difference is regulated by *dsx* and may be important for proper formation of the ovary vs. the testis niche. Rather than being strictly a

genetic process, results from our experiments demonstrate that sexual differentiation in the gonad occurs through a combination of signals that include sex specific hormone signaling. Since the formation of the gonad may represent processes that are conserved from flies to man, this research will provide insight into conserved genes that regulate developmentally similar pathways whose outcome generates major differences observed between the sexes.

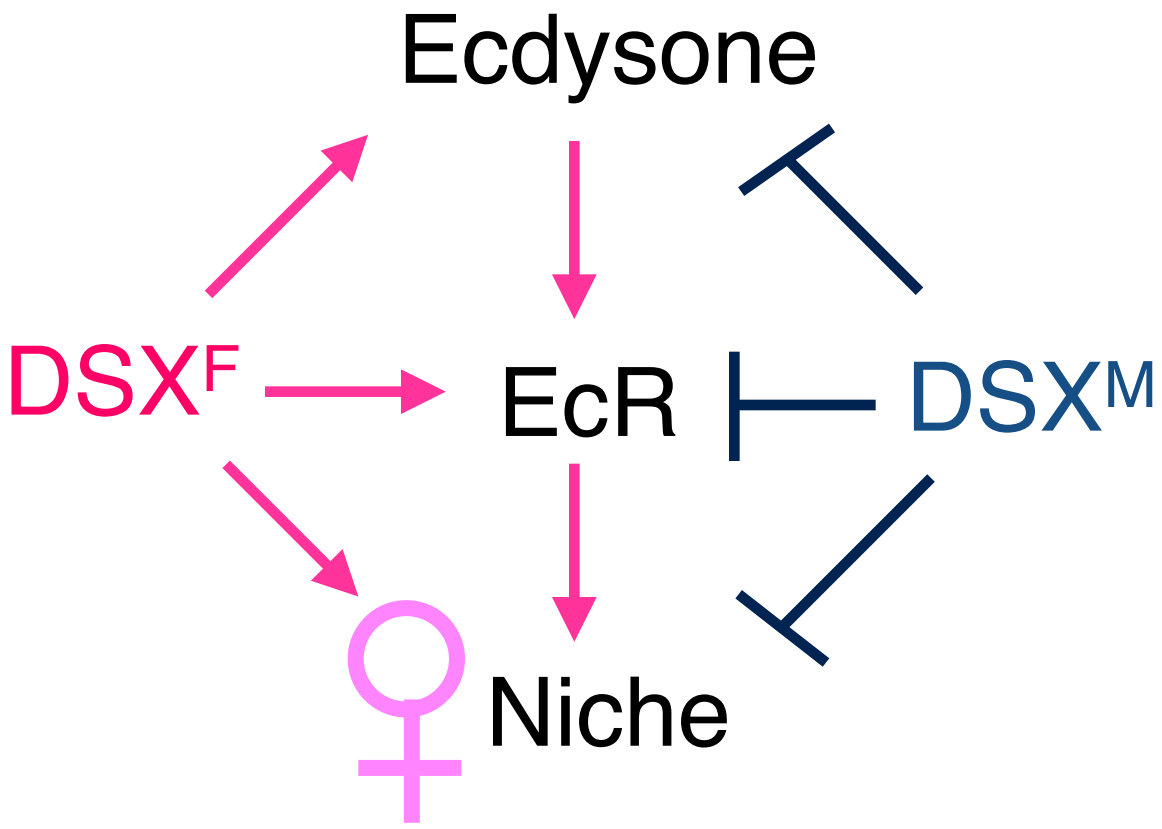


Figure 4.17. Model of the link between hormonal signaling and sex, for sex-specific development of male vs. female germline stem cell niches.

CHAPTER 5: CONCLUSIONS

Sex determination pathways are diverse throughout the animal kingdom, but converge upon conserved genes that encode products that regulate sexual dimorphism. One such conserved factor is represented by the *Drosophila doublesex* (*dsx*) gene, which encodes a sex-specific transcription factor. *dsx* homologs (*dsx*, *mab-3* related transcription factors, DMRTs) play roles in sexual differentiation in a diverse array of metazoans such as, frogs, fish, birds, reptiles, mice, and man.

In *Drosophila*, nearly all manifestations of sexual dimorphism outside the nervous system between males and females are regulated by *doublesex*, yet there are only a few known direct targets of DSX. DSX was identified in 1965 (Hildreth, 1965) and cloned in 1988 (Baker and Wolfner, 1988), but there were still few defined DSX targets and these cannot explain the full array of sexually dimorphic morphologies and behaviors regulated by *dsx* in *D. melanogaster*.

In this work, we sought to identify DSX targets and to understand how DSX contributes to sex- and tissue-specific development and to catalog DSX targets genes in tissues that express *dsx*. In collaboration with a group of scientists from Dr. Brian Oliver, Dr. Stephen Goodwin, and Dr. Teresa Przytycka labs, we determined where DSX is bound in different cell types, which sites are evolutionarily conserved in 20 sequenced *Drosophila* species and in the mouse, the relationship between site strength and occupancy, and which genes respond to acute changes to DSX^F/DSX^M isoform abundance. We then performed RNAi knockdown and dosage-sensitive genetic interaction tests of candidate

targets and found that they resulted in striking tissue-specific transformations of subsets of sexually dimorphic structures.

We found that DSX is bound at largely overlapping sets of genes, regardless of the tissue being analyzed or the DSX isoform. Although we observed DSX binding at thousands of genes in the fat body, and we found that many of these DSX sites are conserved in the Drosophila phylogeny, we observed only a handful of DSX target genes with robust expression changes in the adult fat body when the DSX isoform was acutely switched from male to female, or *vice versa*. Thus, only a few genes were functionally regulated by DSX, despite DSX occupancy, even though both measurements were made at a matched developmental time, place, and experimental condition. We conclude that, in the contexts we examined, DSX binding at a gene confers the possibility of sex-specific regulation; however, functional regulation of a target requires other inputs such as tissue, temporal, spatial, and/or hormonal factors.

This work provides a rich set of DSX target genes for future studies and outlines the mechanisms of DSX action. DSX sex-specific isoforms often bind the same genes, where context-specific factors determine the consequences of that binding. These complex, context-dependent patterns mean that DSX^F can act as a positive regulator of a target gene in one tissue, and DSX^M can act a positive regulator of the same locus in another. DSX acts by a combination of delegating control to transcription factor target genes and by directly micromanaging terminal differentiation genes in a tightly integrated

dance of regulatory inputs. While we may still have decades of research on the roles of DMRT genes in sex determination and differentiation, we now have a comprehensive target gene resource to guide that effort.

While many tissues appear to be regulated by DSX to manifest sexual dimorphism, the *Drosophila* gonad stem cell niche represents an excellent model to dissect how DSX acts on a particular time and place to promote development of a sexually dimorphic tissue. We hypothesize that the presence of a male or female niche in the gonad is regulated autonomously, by *dsx*, and non-autonomously, through cell signaling in the gonad. Using the gonad as a system to study sexual dimorphism, we sought to address how the sex determination pathway and non-autonomous signals create a sexually dimorphic gonad.

The formation of sex-specific structures, such as the hub or terminal filaments, can still form in the complete absence of *dsx* independently of the chromosomal constitution of the animal (XX vs. XY). The hubs and terminal filaments that form in *dsx* mutants have many of the characteristics of the wild-type structures, including the proper morphology and pattern of gene expression, and they associate with and signal to the germ cells normally (Camara, unpublished). DSX does not independently activate the many different genes that are likely to be required to form these structures. Instead, DSX acts to ensure that the male structures form reliably in XY animals, while the female structures form in XX animals. To achieve this, DSX is likely to regulate male- or

female-specific developmental programs that can then function independently of DSX. Thus, DSX would act to "tip the balance" between whether the male (hub) or female (terminal filaments) pathway was activated. One way in which DSX might regulate this balance is by influencing the expression of key upstream regulators of these pathways, such as ecdysone signaling. Using our extensive genomics data in combination with developmental biology of male vs. female germline stem cell niche formation, ecdysone signaling represents the best candidate for a "pro female gene" in our model of sex-specific gonad niche development.

Our finding that ecdysone signaling activity in the larval female gonad, but not the male indicates that the gonads have dimorphic components involved in activation of ecdysone signaling. Differences in how, where, and when sex-specific gonads respond to ecdysone signaling may be attributed to male and female gonads carrying out different functions. Even though other tissues in the male must respond to ecdysone to control metamorphosis, the testis remains much as it has since the end of embryogenesis with continued spermatogenesis.

Ecdysteroid levels in adult males have been shown to be lower compared to the high-level pulses produced by the prothoracic gland that drive transitions during development and those of adult female insects. In male adults, genes important for the biosynthesis of ecdysone are lower in expression compared to larvae, pupae, and females (Hentze et al., 2013). Indeed, when we examined the spatial patterns of ecdysone ligand

sensor activation in male and female gonads at the onset of metamorphosis, we detected activation in the somatic cells of the ovary at higher levels compared to the testis. Thus, we show here that male and female gonads respond differently to the late larval ecdysteroid pulse. In support of differential activity of ecdysone signaling in male vs. female larval gonads, we have found that males and females exhibit a dramatically different response to ecdysone in the gonads. By examining a reporter for ecdysone receptor activity (ecdysone receptor DNA binding sites upstream of a lacZ reporter, EcRE-lacZ), we find that the larval ovary shows robust activity of this reporter while the same stage larval testis shows little activity.

The most straightforward way to have female-specific ecdysone response would be to have the ecdysone receptor (*EcR*) expressed only in the female and not in the male gonad. Indeed, we find EcR in all somatic nuclei of the larval ovary, but not the testis. We hypothesize that the sex-differences we observe in ecdysone signaling is important for how *dsx* regulates the male and female paths in the gonads. If *dsx* modulates the ecdysone response so that it is robustly on in females, to promote ovary formation, and off in males, to allow for testis maintenance, then we expect in *dsx* mutants for there to be an intermediate ecdysone response in the gonads of both sexes (XX and XY). In support of this model we detect EcR in all somatic nuclei of *dsx* mutant gonads with a male niche component, the hub. Using reporter constructs driven by EcR enhancer elements with putative DSX binding sites, we show that there is at least one region of the EcR locus that

is expressed in a tissue-specific manner and that this expression in the gonad is regulated by DSX. We show that knockdown of EcR not only inhibits terminal filament formation, but can promote hub formation. We also show that over-expression of EcR, can repress the hub from forming. These data suggest that the male form of *dsx*, dsx^M, is important for repressing EcR expression and ecdysone signaling in the testis. Whereas, the female form of *dsx*, dsx^F, is important for ensuring female structures form in XX animals. Thus it appears that *Ecdysone receptor* acts as a "pro female gene" and that *dsx* allows the steroid hormone ecdysone to act differently in males and females, so that males ignore this signal and maintain the hub and testis, while females produce a robust response to form ovaries.

Rather than being strictly a genetic process, results from our experiments demonstrate that sexual differentiation in the gonad occurs through a combination of signals that include sex specific hormone signaling. Since the formation of the gonad may represent processes that are conserved from flies to man, this research will provide insight into conserved genes that regulate developmentally similar pathways whose outcome generates major differences observed between the sexes.

REFERENCES

Adam, G., Perrimon, N., and Noselli, S. (2003). The retinoic-like juvenile hormone controls the looping of left-right asymmetric organs in Drosophila. Development (Cambridge, England) 130, 2397-2406.

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., et al. (2000). The genome sequence of Drosophila melanogaster. Science (New York, NY) 287, 2185-2195.

Ahmad, S. (1979). The functional roles of cytochrome P-450 mediated systems: Present knowledge and future areas of investigations. Drug Metabolism Reviews, 10(1), 1-14. doi: 10.3109/03602537908993898

Ahmad, S.M., and Baker, B.S. (2002). Sex-specific deployment of FGF signaling in Drosophila recruits mesodermal cells into the male genital imaginal disc. Cell 109, 651-661.

Anand, A., Villella, A., Ryner, L. C., Carlo, T., Goodwin, S. F., Song, H. J., Taylor, B. J. (2001). Molecular genetic dissection of the sex-specific and vital functions of the drosophila melanogaster sex determination gene fruitless. Genetics, 158(4), 1569-1595.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome biology 11, R106.

Anders, S., Pyl, P.T., and Huber, W. (2014). HTSeq – A Python framework to work with high-throughput sequencing data. bioRxiv preprint.

Arbeitman, M. N. et al. (2004). A genomic analysis of Drosophila somatic sexual

Arbeitman, M. N., Fleming, A. A., Siegal, M. L., Null, B. H., & Baker, B. S. (2004). A genomic analysis of drosophila somatic sexual differentiation and its regulation. Development (Cambridge, England), 131(9), 2007-2021. doi:10.1242/dev.01077

Auer, P. L., and Doerge, R.W. (2010). Statistical design and analysis of RNA sequencing Data. Genetics, 185:405-416.

Baker, B. S., & Wolfner, M. F. (1988). A molecular analysis of doublesex, a bifunctional gene that controls both male and female sexual differentiation in drosophila melanogaster. Genes & Development, 2(4), 477-489.

Baker, B. S., Burtis, K., Goralski, T., Mattox, W., & Nagoshi, R. (1989). Molecular genetic aspects of sex determination in drosophila melanogaster. Genome / National Research Council Canada = Genome / Conseil National De Recherches Canada, 31(2), 638-645.

Baker, B.S., and Wolfner, M.F. (1988). A molecular analysis of doublesex, a bifunctional gene that controls both male and female sexual differentiation in Drosophila melanogaster. Genes & development 2, 477-489.

Bardot, O., Godt, D., Laski, F.A., Courderc, J.L. (2002). Expressing UAS-bab1 and UAS-bab2: a comparative study of gain-of-function effects and the potential to rescue the bric-a-brac mutant phenotype. Genesis. 34, 66-70.

Barmina, O., and Kopp, A. (2007). Sex-specific expression of a HOX gene associated with rapid morphological evolution. Dev Biol. 311, 277-86.

Barmina, O., Gonzalo, M., McIntyre, L.M., and Kopp, A. (2005). Sex- and segment-specific modulation of gene expression profiles in Drosophila. Developmental biology 288, 528-544.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., et al. (2013). NCBI GEO: archive for functional genomics data sets--update. Nucleic acids research 41, D991-995.

Bayrer, J.R., Zhang, W., and Weiss, M.A. (2005). Dimerization of doublesex is mediated by a cryptic ubiquitin-associated domain fold: implications for sex-specific gene regulation. The Journal of biological chemistry 280, 32989-32996.

Belote, J. M., Handler, A. M., Wolfner, M. F., Livak, K. J., Baker, B. S. (1985). Sex-specific regulation of yolk protein gene expression in Drosophila. Cell. 40, 339-348.

Belote, J. M., McKeown, M., Boggs, R. T., Ohkawa, R., & Sosnowski, B. A. (1989). Molecular genetics of transformer, a genetic switch controlling sexual differentiation in drosophila. Developmental Genetics, 10(3), 143-154. doi:10.1002/dvg.1020100304

Biswas, D., Milne, T.A., Basrur, V., Kim, J., Elenitoba-Johnson, K.S., Allis, C.D., and Roeder, R.G. (2011). Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. Proceedings of the National Academy of Sciences of the United States of America 108, 15751-15756.

Bitoun, E., Oliver, P.L., and Davies, K.E. (2007). The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. Human molecular genetics 16, 92-106.

Boggs, R., Gregor, P., Idriss, S., Belote, J. M., McKeown, M. (1987). Regulation of sexual differentiation in D. melanogaster via alternative splicing of RNA from the transformer gene. Cell. 50, 739-747.

Bopp, D., Horabin, J. I., Lersch, R. A., Cline, T. W., & Schedl, P. (1993). Expression of the sex-lethal gene is controlled at multiple levels during drosophila oogenesis. Development (Cambridge, England), 118(3), 797-812.

Bownes, M., Ronaldson, E., and Mauchline, D. (1996). 20-Hydroxyecdysone, but not juvenile hormone, regulation of yolk protein gene expression can be mapped to cis-acting DNA sequences. Developmental biology 173, 475-489.

Brennan, C. A., Li, T. R., Bender, M., Hsiung, F., & Moses, K. (2001). Broad-complex, but not ecdysone receptor, is required for progression of the morphogenetic furrow in the drosophila eye. Development (Cambridge, England), 128(1), 1-11.

Burtis, K. C., & Baker, B. S. (1989). Drosophila doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. Cell, 56(6), 997-1010. doi:0092-8674(89)90633-8 [pii]

Burtis, K.C., and Baker, B.S. (1989). Drosophila doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. Cell 56, 997-1010.

Burtis, K.C., Coschigano, K.T., Baker, B.S., and Wensink, P.C. (1991). The doublesex proteins of Drosophila melanogaster bind directly to a sex-specific yolk protein gene enhancer. The EMBO journal 10, 2577-2582.

Cabrera, G., et al. (2002). Expression pattern of Gal4 enhancer trap insertions into the bric à brac locus generated by P-element replacement. Genesis. 34, 62-65.

Camara, N., Whitworth, C., & Van Doren, M. (2008). The creation of sexual dimorphism in the drosophila soma. Current Topics in Developmental Biology, 83, 65-107. doi:10.1016/S0070-2153(08)00403-1

Casper, A.L. and M. Van Doren, The establishment of sexual identity in the Drosophila germline. Development, 2009. 136(22): p. 3821-30.

Chang, P. L., Dunham, J. P., Nuzhdin, S. V., & Arbeitman, M. N. (2011). Somatic sex-specific transcriptome differences in drosophila revealed by whole transcriptome sequencing. BMC Genomics, 12, 364-2164-12-364. doi:10.1186/1471-2164-12-364

Chatterjee, S.S., Uppendahl, L.D., Chowdhury, M.A., Ip, P.L., and Siegal, M.L. (2011). The female-specific doublesex isoform regulates pleiotropic transcription factors to pattern genital development in Drosophila. Development (Cambridge, England) 138, 1099-1109.

Chen, E. and Baker, B.S. (1997). Compartmental organization of the Drosophila genital imaginal discs. Development. 124, 205-218.

Chen, Z.X., Sturgill, D., Qu, J., Jiang, H., Park, S., Boley, N., Suzuki, A.M., Fletcher, A.R., Plachetzki, D.C., FitzGerald, P.C., et al. (2014). Comparative validation of the D. melanogaster modENCODE transcriptome annotation. Genome research 24, 1209-1223.

Cherbas, L., Lee, K., & Cherbas, P. (1991). Identification of ecdysone response elements by analysis of the drosophila Eip28/29 gene. Genes & Development, 5(1), 120-131.

Chertemps, T., Duportets, L., Labeur, C., Ueyama, M., and Wicker-Thomas, C. (2006). A female-specific desaturase gene responsible for diene hydrocarbon biosynthesis and courtship behaviour in Drosophila melanogaster. Insect Mol Biol 15, 465-473.

Christiansen, A., et al. (2002). Sex comes in from the cold: the integration of sex and pattern. Trends Genet. 18, 510-516.

Clark, I.B., A.P. Jarman, and D.J. Finnegan, Live imaging of Drosophila gonad formation reveals roles for Six4 in regulating germline and somatic cell migration. BMC Dev Biol, 2007. 7: p. 52.

Cline, T. W. (1986). A female-specific lethal lesion in an X-linked positive regulator of the drosophila sex determination gene, sex-lethal. Genetics, 113(3), 641-663.

Cock, P.J., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics (Oxford, England) 25, 1422-1423.

Coschigano, K., and Wensink, P. (1993). Sex-specific transcriptional regulation by the male and female doublesex proteins of Drosophila. Genes Dev. 7, 42-54.

Couderc, J., et al. (2002). The bric à brac locus consists of two paralogous genes encoding BTB/POZ domain proteins and acts as a homeotic and morphogenetic regulator of imaginal development in Drosophila. Development. 129, 2419-2433.

Dalton, J. E., Lebo, M. S., Sanders, L. E., Sun, F., & Arbeitman, M. N. (2009). Ecdysone receptor acts in fruitless- expressing neurons to mediate drosophila courtship behaviors. Current Biology : CB, 19(17), 1447-1452. doi:10.1016/j.cub.2009.06.063

Davies, E., and Fuller, M. (2008). Regulation of self-renewal and differentiation in adult stem cell lineages; lessons from the Drosophila male germ line. Cold Spring Harb Symp Quant Biol. 73, 137-145.

DE LOOF, A. (2006), Ecdysteroids: the overlooked sex steroids of insects? Males: the black box. Insect Science, 13: 325–338. doi: 10.1111/j.1744-7917.2006.00101.x

De Loof, A., & Huybrechts, R. (1998). Insects do not have sex hormones: A myth? General and Comparative Endocrinology, 111(3), 245-260. doi:S0016-6480(98)97101-5 [pii]

DeFalco, T., Camara, N., Le Bras, S., and Van Doren, M. (2008). Nonautonomous sex determination controls sexually dimorphic development of the Drosophila gonad. Developmental cell 14, 275-286.

DeFalco, T.J., G. Verney, A.B. Jenkins, J.M. McCaffery, S. Russell, and M. Van Doren, Sex-specific apoptosis regulates sexual dimorphism in the Drosophila embryonic gonad. Dev Cell, 2003. 5(2): p. 205-216.

Devi, T.R., and Shyamala, B.V. (2013). Male- and female-specific variants of doublesex gene products have different roles to play towards regulation of Sex combs reduced expression and sex comb morphogenesis in Drosophila. Journal of biosciences 38, 455-460.

Drosophila 12 Genomes, C., Clark, A.G., Eisen, M.B., Smith, D.R., Bergman, C.M., Oliver, B., Markow, T.A., Kaufman, T.C., Kellis, M., Gelbart, W., et al. (2007). Evolution of genes and genomes on the Drosophila phylogeny. Nature 450, 203-218.

Eggers, S., and Sinclair, A. (2012). Mammalian sex determination-insights from humans and mice. Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology 20, 215-238.

Erdman, S.E., Chen, H.J., and Burtis, K.C. (1996). Functional and genetic characterization of the oligomerization and DNA binding properties of the Drosophila doublesex proteins. Genetics 144, 1639-1652.

Erickson, J. W., & Cline, T. W. (1993). A bZIP protein, sisterless-a, collaborates with bHLH transcription factors early in drosophila development to determine sex. Genes & Development, 7(9), 1688-1702.

Estrada, B., Casares, F., and Sanchez-Herrero, E. (2003). Development of the genitalia in Drosophila melanogaster. Differentiation; research in biological diversity 71, 299-310.

Feng, Q., Wang, H., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Struhl, K., and Zhang, Y. (2002). Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. Current biology : CB 12, 1052-1058.

Fisher, W.W., Li, J.J., Hammonds, A.S., Brown, J.B., Pfeiffer, B.D., Weiszmann, R., MacArthur, S., Thomas, S., Stamatoyannopoulos, J.A., Eisen, M.B., et al. (2012). DNA regions bound at low occupancy by transcription factors do not drive patterned reporter gene expression in Drosophila. Proceedings of the National Academy of Sciences of the United States of America 109, 21330-21335.

Flicek, P., Ahmed, I., Amode, M.R., Barrell, D., Beal, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fairley, S., et al. (2013). Ensembl 2013. Nucleic acids research 41, D48-55.

Foronda, D., Martin, P., and Sanchez-Herrero, E. (2012). Drosophila Hox and sex-determination genes control segment elimination through EGFR and extramacrochetae activity. PLoS genetics 8, e1002874.

Gan, Q., Chepelev, I., Wei, G., Tarayrah, L., Cui, K., Zhao, K., & Chen, X. (2010). Dynamic regulation of alternative splicing and chromatin structure in drosophila gonads revealed by RNA-seq. Cell Research, 20(7), 763-783. doi:10.1038/cr.2010.64

Gancz, D., & Gilboa, L. (2013). Insulin and target of rapamycin signaling orchestrate the development of ovarian niche-stem cell units in drosophila. Development (Cambridge, England), 140(20), 4145-4154. doi:10.1242/dev.093773

Gancz, D., Lengil, T., & Gilboa, L. (2011). Coordinated regulation of niche and stem cell precursors by hormonal signaling. PLoS Biology, 9(11), e1001202. doi:10.1371/journal.pbio.1001202

Garrett-Engele, C.M., Siegal, M.L., Manoli, D.S., Williams, B.C., Li, H., and Baker, B.S. (2002). intersex, a gene required for female sexual development in Drosophila, is expressed in both sexes and functions together with doublesex to regulate terminal differentiation. Development (Cambridge, England) 129, 4661-4675.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome biology 5, R80.

Gilboa, L. and Lehmann, R. (2006). Soma-germline interactions coordinate homeostasis and growth in the Drosophila gonad. 443, 97-100.

Godt, D. and Laski, F. (1995). Mechanisms of cell rearrangement and cell recruitment in Drosophila ovary morphogenesis and the requirement of bric à brac. Development. 121, 173-187.

Gonczy, P., Viswanathan, S., & DiNardo, S. (1992). Probing spermatogenesis in drosophila with P-element enhancer detectors. Development (Cambridge, England), 114(1), 89-98.

Goralski, T. J., Edstrom, J. E., & Baker, B. S. (1989). The sex determination locus transformer-2 of drosophila encodes a polypeptide with similarity to RNA binding proteins. Cell, 56(6), 1011-1018. doi:0092-8674(89)90634-X [pii]

Gorfinkiel, N., Sanchez, L., and Guerrero, I. (2003). Development of the Drosophila genital disc requires interactions between its segmental primordia. Development (Cambridge, England) 130, 295-305.

Greil, F., Moorman, C., and van Steensel, B. (2006). DamID: mapping of in vivo protein-genome interactions using tethered DNA adenine methyltransferase. Methods in enzymology 410, 342-359.

Gupta, S., Stamatoyannopoulos, J.A., Bailey, T.L., and Noble, W.S. (2007). Quantifying similarity between motifs. Genome biology 8, R24.

Hall, B. L., & Thummel, C. S. (1998). The RXR homolog ultraspiracle is an essential component of the drosophila ecdysone receptor. Development (Cambridge, England), 125(23), 4709-4717.

Hardy RW, Tokuyasu KT, Lindsley DL, Garavito M. The germinal proliferation center in the testis of Drosophila melanogaster. J Ultrastruct Res 1979; 69:180 - 90; http://dx.doi.org/10.1016/S0022-5320(79)90108-4; PMID: 114676

Hempel, L. and Oliver, B. (2007). Sex-specific DoublesexM expression in subsets of Drosophila somatic gonad cells. BMC Dev Biol. 7, 113.

Henrich, V. C., Szekely, A. A., Kim, S. J., Brown, N. E., Antoniewski, C., Hayden, M. A.,Gilbert, L. I. (1994). Expression and function of the ultraspiracle (usp) gene during development of drosophila melanogaster. Developmental Biology, 165(1), 38-52. doi:S0012-1606(84)71232-2

Hentze, J. L., Moeller, M. E., Jorgensen, A. F., Bengtsson, M. S., Bordoy, A. M., Warren, J. T., Rewitz, K. F. (2013). Accessory gland as a site for prothoracicotropic hormone controlled ecdysone synthesis in adult male insects. PloS One, 8(2), e55131. doi:10.1371/journal.pone.0055131

Hertel, K., et al. (1996). Structural and functional conservation of the Drosophila doublesex splicing enhancer repeat elements. RNA. 2, 969-981.

Hildreth, P. (1965). Doublesex, recessive gene that transforms both males and females of Drosophila ino intersexes. Genetics. 51, 659-678.

Hildreth, P.E. (1965). Doublesex, Recessive Gene That Transforms Both Males and Females of Drosophila into Intersexes. Genetics 51, 659-678.

Hodin, J., & Riddiford, L. M. (1998). The ecdysone receptor and ultraspiracle regulate the timing and progression of ovarian morphogenesis during drosophila metamorphosis. Development Genes and Evolution, 208(6), 304-317.

Jakob, S. and Lovell-Badge, R. (2011) Sex determination and the control of Sox9 expression in mammals. FEBS J. 278(7): p. 1002-9.

Jemc, J.C. (2011). Somatic gonadal cells: the supporting cast for the germline. Genesis. 49(10): p. 753-75.

Jeong, S., et al. (2006). Regulation of body pigmentation by the Abdominal-B Hox protein and its gain and loss in Drosophila evolution. Cell. 125, 1387-1399.

Jursnich, V.A., and Burtis, K.C. (1993). A positive role in differentiation for the male doublesex protein of Drosophila. Developmental biology 155, 235-249.

Keisman, E.L., Christiansen, A.E., and Baker, B.S. (2001). The sex determination gene doublesex regulates the A/P organizer to direct sex-specific patterns of growth in the Drosophila genital imaginal disc. Developmental cell 1, 215-225.

Kharchenko, P.V., Tolstorukov, M.Y., and Park, P.J. (2008). Design and analysis of ChIP-seq experiments for DNA-binding proteins. Nature biotechnology 26, 1351-1359.

Kiger, A., Cooper, H., Fuller, M. (2000). Somatic support cells restrict germline stem cell self-renewal and promote differentiation. Nature. 407, 750-754.

Kim, S., Bardwell, V. J., & Zarkower, D. (2007). Cell type-autonomous and non-autonomous requirements for Dmrt1 in postnatal testis differentiation. Developmental Biology, 307(2), 314-327. doi:S0012-1606(07)00869-X [pii]

Kim, Y., and Capel, B. (2006). Balancing the bi-potential gonad between alternative organ fates: a new perspective on an old problem. Developmental Dynamics. 235, 2292-2300.

King-Jones, K., Charles, J. P., Lam, G., & Thummel, C. S. (2005). The ecdysone-induced DHR4 orphan nuclear receptor coordinates growth and maturation in drosophila. Cell, 121(5), 773-784. doi:S0092-8674(05)00340-5 [pii]

Kitadate, Y., and Kobayashi, S. (2010). Notch and Egfr signaling act antagonistically to regulate germ line stem cell niche formation in Drosophila male embryonic gonads. PNAS. 107, 14241-14246.

Kitadate, Y., et al. (2007). Boss/Sev signaling from germline to soma restricts germline stem cell niche formation in the anterior region of Drosophila male gonads. Dev Cell. 13, 151-159.

Koelle, M. R., Talbot, W. S., Segraves, W. A., Bender, M. T., Cherbas, P., & Hogness, D. S. (1991). The drosophila EcR gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. Cell, 67(1), 59-77. doi:0092-8674(91)90572-G [pii]

Kopp, A., et al. (2000). Genetic control and evolution of sexually dimorphic characters in Drosophila. Nature. 408, 553-559.

Kozlova, T., & Thummel, C. S. (2002). Spatial patterns of ecdysteroid receptor activation during the onset of drosophila metamorphosis. Development (Cambridge, England), 129(7), 1739-1750.

Kozlova, T., & Thummel, C. S. (2003). Essential roles for ecdysone signaling during drosophila mid-embryonic development. Science (New York, N.Y.), 301(5641), 1911-1914. doi:10.1126/science.1087419

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology 10, R25.

Le Bras, S., & Van Doren, M. (2006). Development of the male germline stem cell niche in drosophila. Developmental Biology, 294(1), 92-103. doi:S0012-1606(06)00127-8 [pii]

Lebo, M.S., Sanders, L.E., Sun, F., and Arbeitman, M.N. (2009). Somatic, germline and sex hierarchy regulated gene expression during Drosophila metamorphosis. BMC genomics 10, 80.

Lee, G., Hall, J.C., and Park, J.H. (2002). Doublesex gene expression in the central nervous system of Drosophila melanogaster. J Neurogenet 16, 229-248.

Leonie Quinn, Jane Lin, Nicola Cranna, Jue Er Amanda Lee, Naomi Mitchell and Ross Hannan (2012). Steroid Hormones in Drosophila: How Ecdysone Coordinates Developmental Signalling with Cell Growth and Division, Steroids - Basic Science, Prof. Hassan Abduljabbar (Ed.), ISBN: 978-953-307-866-3, InTech, DOI: 10.5772/27927. Available from: http://www.intechopen.com/books/steroids-basic-science/steroid-hormones-in-drosophila-how-ecdysone-coordinates-developmental-signalling-with-cell-growth-an

Lours, C., et al. (2003). The Drosophila melanogaster BTB proteins bric à brac bind DNA through a composite DNA binding domain containing a pipsqueak and an AT-Hook motif. Nucleic Acids Res 31, 5389-5398.

Luo, S., Shi, G. and Baker, B. (2011). Direct targets of the D. melanogaster DSXF protein and the evolution of sexual development. Development. 138, 2761-2771.

Machanick, P., and Bailey, T.L. (2011). MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics (Oxford, England) 27, 1696-1697.

Mahowald, A. P., & Wei, G. (1994). Sex determination of germ cells in drosophila. Ciba Foundation Symposium, 182, 193-202; discussion 202-9.

Markopoulou, K., & Artavanis-Tsakonas, S. (1991). Developmental analysis of the facets, a group of intronic mutations at the notch locus of drosophila melanogaster that affect postembryonic development. The Journal of Experimental Zoology, 257(3), 314-329. doi: 10.1002/jez.1402570305

Mathelier, A., Zhao, X., Zhang, A.W., Parcy, F., Worsley-Hunt, R., Arenillas, D.J., Buchman, S., Chen, C.Y., Chou, A., Ienasescu, H., et al. (2013). JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic acids research.

Mattox, W., McGuffin, M. E., & Baker, B. S. (1996). A negative feedback mechanism revealed by functional analysis of the alternative isoforms of the drosophila splicing regulator transformer-2. Genetics, 143(1), 303-314.

Mattox, W., Palmer, M. J., & Baker, B. S. (1990). Alternative splicing of the sex determination gene transformer-2 is sex-specific in the germ line but not in the soma. Genes & Development, 4(5), 789-805.

McKeown, M., Belote, J. M., & Baker, B. S. (1987). A molecular analysis of transformer, a gene in drosophila melanogaster that controls female sexual differentiation. Cell, 48(3), 489-499. doi: 0092-8674(87)90199-1 [pii]

McKeown, M., et al. (1988). Ectopic expression of the female transformer gene product leads to female differentiation of chromosomally male Drosophila. Cell. 53, 887-895.

Moore, L.A., H.T. Broihier, M. Van Doren, and R. Lehmann, Gonadal mesoderm and fat body initially follow a common developmental path in Drosophila. Development, 1998. 125(5): p. 837-44.

Morris, L. X., & Spradling, A. C. (2012). Steroid signaling within drosophila ovarian epithelial cells sex-specifically modulates early germ cell development and meiotic entry. PloS One, 7(10), e46109. doi:10.1371/journal.pone.0046109

Mueller, D., Bach, C., Zeisig, D., Garcia-Cuellar, M.P., Monroe, S., Sreekumar, A., Zhou, R., Nesvizhskii, A., Chinnaiyan, A., Hess, J.L., et al. (2007). A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. Blood 110, 4445-4454.

Mugat, B., Brodu, V., Kejzlarova-Lepesant, J., Antoniewski, C., Bayer, C. A., Fristrom, J. W., & Lepesant, J. A. (2000). Dynamic expression of broad-complex isoforms mediates temporal control of an ecdysteroid target gene at the onset of drosophila metamorphosis. Developmental Biology, 227(1), 104-117. doi:10.1006/dbio.2000.9879

Murphy, M.W., Sarver, A.L., Rice, D., Hatzi, K., Ye, K., Melnick, A., Heckert, L.L., Zarkower, D., and Bardwell, V.J. (2010). Genome-wide analysis of DNA binding and transcriptional regulation by the mammalian Doublesex homolog DMRT1 in the juvenile testis. Proceedings of the National Academy of Sciences of the United States of America 107, 13360-13365.

Murphy, M.W., Zarkower, D., and Bardwell, V.J. (2007). Vertebrate DM domain proteins bind similar DNA sequences and can heterodimerize on DNA. BMC molecular biology 8, 58.

Nagoshi, R. N., & Baker, B. S. (1990). Regulation of sex-specific RNA splicing at the drosophila doublesex gene: Cis-acting mutations in exon sequences alter sex-specific RNA splicing patterns. Genes & Development, 4(1), 89-97.

Nagoshi, R. N., Patton, J. S., Bae, E., & Geyer, P. K. (1995). The somatic sex determines the requirement for ovarian tumor gene activity in the proliferation of the drosophila germline. Development (Cambridge, England), 121(2), 579-587.

Nagoshi, R.N., McKeown, M., Burtis, K.C., Belote, J.M., and Baker, B.S. (1988). The control of alternative splicing at genes regulating sexual differentiation in D. melanogaster. Cell 53, 229-236.

Nanda, S., T.J. DeFalco, S.H. Loh, N. Phochanukul, N. Camara, M. Van Doren, and S. Russell, Sox100B, a Drosophila group E Sox-domain gene, is required for somatic testis differentiation. Sex Dev, 2009. 3(1): p. 26-37.

Negre, N., Brown, C.D., Ma, L., Bristow, C.A., Miller, S.W., Wagner, U., Kheradpour, P., Eaton, M.L., Loriaux, P., Sealfon, R., et al. (2011). A cis-regulatory map of the Drosophila genome. Nature 471, 527-531.

Negri, I., Pellecchia, M., Greve, P., Daffonchio, D., Bandi, C., & Alma, A. (2010). Sex and stripping: The key to the intimate relationship between wolbachia and host? Communicative & Integrative Biology, 3(2), 110-115.

Neville, M.C., Nojima, T., Ashley, E., Parker, D.J., Walker, J., Southall, T., Van de Sande, B., Marques, A.C., Fischer, B., Brand, A.H., et al. (2014). Male-specific fruitless isoforms target neurodevelopmental genes to specify a sexually dimorphic nervous system. Current biology : CB 24, 229-241.

Nguyen, A.T., and Zhang, Y. (2011). The diverse functions of Dot1 and H3K79 methylation. Genes & development 25, 1345-1358.

Nguyen, A.T., Xiao, B., Neppl, R.L., Kallin, E.M., Li, J., Chen, T., Wang, D.Z., Xiao, X., and Zhang, Y. (2011). DOT1L regulates dystrophin expression and is critical for cardiac function. Genes & development 25, 263-274.

Okada, Y., Feng, Q., Lin, Y., Jiang, Q., Li, Y., Coffield, V.M., Su, L., Xu, G., and Zhang, Y. (2005). hDOT1L links histone methylation to leukemogenesis. Cell 121, 167-178.

Oliver, B., Kim, Y.J., and Baker, B.S. (1993). Sex-lethal, master and slave: a hierarchy of germ-line sex determination in Drosophila. Development (Cambridge, England) 119, 897-908.

Palanker, L., Necakov, A. S., Sampson, H. M., Ni, R., Hu, C., Thummel, C. S., & Krause, H. M. (2006). Dynamic regulation of drosophila nuclear receptor activity in vivo. Development (Cambridge, England), 133(18), 3549-3562. doi:dev.02512 [pii]

Parisi, M.J., Gupta, V., Sturgill, D., Warren, J.T., Jallon, J.M., Malone, J.H., Zhang, Y., Gilbert, L.I., and Oliver, B. (2010). Germline-dependent gene expression in distant non-gonadal somatic tissues of Drosophila. BMC genomics 11, 346.

Pitman, J. L., Tsai, C. C., Edeen, P. T., Finley, K. D., Evans, R. M., & McKeown, M. (2002). DSF nuclear receptor acts as a repressor in culture and in vivo. Developmental Biology, 245(2), 315-328. doi:10.1006/dbio.2002.0648

Randsholt, N. and Santamaria, P. (2008). How Drosophila change their combs: the Hox gene Sex combs reduced and sex comb variation among Sophophora species. Evol Dev. 10, 121-133.

Raymond, C., et al. (1998). Evidence for evolutionary conservation of sex-determining genes. Nature. 391, 691-695.

Raymond, C.S., Parker, E.D., Kettlewell, J.R., Brown, L.G., Page, D.C., Kusz, K., Jaruzelska, J., Reinberg, Y., Flejter, W.L., Bardwell, V.J., et al. (1999). A region of human chromosome 9p required for testis development contains two genes related to known sexual regulators. Human molecular genetics 8, 989-996.

Richards, S., Liu, Y., Bettencourt, B.R., Hradecky, P., Letovsky, S., Nielsen, R., Thornton, K., Hubisz, M.J., Chen, R., Meisel, R.P., et al. (2005). Comparative genome sequencing of Drosophila pseudoobscura: chromosomal, gene, and cis-element evolution. Genome research 15, 1-18.

Rideout, E.J., Dornan, A.J., Neville, M.C., Eadie, S., and Goodwin, S.F. (2010). Control of sexual differentiation and behavior by the doublesex gene in Drosophila melanogaster. Nature neuroscience 13, 458-466.

Riechmann, V., K.P. Rehorn, R. Reuter, and M. Leptin, The genetic control of the distinction between fat body and gonadal mesoderm in Drosophila. Development, 1998. 125(4): p. 713-23.

Robinett, C. C., Vaughan, A. G., Knapp, J. M., & Baker, B. S. (2010). Sex and the single cell. II. there is a time and place for sex. PLoS Biology, 8(5), e1000365. doi:10.1371/journal.pbio.1000365

Ronaldson, E., and Bownes, M. (1995). Two independent cis-acting elements regulate the sex- and tissue-specific expression of yp3 in Drosophila melanogaster. Genet Res 66, 9-17.

Ryner, L., and Baker, B. (1991). Regulation of doublesex pre-mRNA processing occurs by 3'-splice site activation. Genes Dev. 5, 2071-2085.

Samson, M.L., and Rabinow, L. (2013). Transcriptomic Analysis of Sexual Differentiation in Somatic Tissues of Drosophila melanogaster: Successes and Caveats. Sexual development : genetics, molecular biology, evolution, endocrinology, embryology, and pathology of sex determination and differentiation.

Sánchez, L. and Guerrero, I. (2001). The development of the Drosophila genital disc. Bioessays. 23, 698-707.

Sanders, L., Arbeitman, M. Doublesex establishes sexual dimorphism in the Drosophila central nervous system in an isoforms dependent manner by directing cell number. Dev Biol. 320, 378-390.

Schwedes, C., Tulsiani, S., & Carney, G. E. (2011). Ecdysone receptor expression and activity in adult drosophila melanogaster. Journal of Insect Physiology, 57(7), 899-907. doi:10.1016/j.jinsphys.2011.03.027

Shanower, G.A., Muller, M., Blanton, J.L., Honti, V., Gyurkovics, H., and Schedl, P. (2005). Characterization of the grappa gene, the Drosophila histone H3 lysine 79 methyltransferase. Genetics 169, 173-184.

Sheng, X. R., Brawley, C. M., & Matunis, E. L. (2009). Dedifferentiating spermatogonia outcompete somatic stem cells for niche occupancy in the drosophila testis. Cell Stem Cell, 5(2), 191-203. doi:10.1016/j.stem.2009.05.024

Shirangi, T.R., Dufour, H.D., Williams, T.M., and Carroll, S.B. (2009). Rapid evolution of sex pheromone-producing enzyme expression in Drosophila. PLoS biology 7, e1000168.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome research 15, 1034-1050.

Siera, S.G., and Cline, T.W. (2008). Sexual back talk with evolutionary implications: stimulation of the Drosophila sex-determination gene sex-lethal by its target transformer. Genetics 180, 1963-1981.

Southall, T.D., and Brand, A.H. (2007). Chromatin profiling in model organisms. Briefings in functional genomics & proteomics 6, 133-140.

Steger, D.J., Lefterova, M.I., Ying, L., Stonestrom, A.J., Schupp, M., Zhuo, D., Vakoc, A.L., Kim, J.E., Chen, J., Lazar, M.A., et al. (2008). DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. Molecular and cellular biology 28, 2825-2839.

Sturgill, D., Malone, J.H., Sun, X., Smith, H.E., Rabinow, L., Samson, M.L., and Oliver, B. (2013). Design of RNA splicing analysis null models for post hoc filtering of Drosophila head RNA-Seq data with the splicing analysis kit (Spanki). BMC bioinformatics 14, 320.

Talbot, W. S., Swyryd, E. A., & Hogness, D. S. (1993). Drosophila tissues with different metamorphic responses to ecdysone express different ecdysone receptor isoforms. Cell, 73(7), 1323-1337. doi:0092-8674(93)90359-X [pii]

Tanaka, K., Barmina, O., Sanders, L.E., Arbeitman, M.N., and Kopp, A. (2011). Evolution of sex-specific traits through changes in HOX-dependent doublesex expression. PLoS biology 9, e1001131.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics (Oxford, England) 25, 1105-1111.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols 7, 562-578.

Villar, D., Flicek, P., and Odom, D.T. (2014). Evolution of transcription factor binding in metazoans - mechanisms and functional implications. Nature reviews Genetics 15, 221-233.

Wang, W., and Yoder, J.H. (2012). Hox-mediated regulation of doublesex sculpts sex-specific abdomen morphology in Drosophila. Developmental dynamics : an official publication of the American Association of Anatomists 241, 1076-1090.

Wang, W., Kidd, B.J., Carroll, S.B., and Yoder, J.H. (2011). Sexually dimorphic regulation of the Wingless morphogen controls sex-specific segment number in Drosophila. Proceedings of the National Academy of Sciences of the United States of America 108, 11139-11144.

Waterbury, J. A., Horabin, J. I., Bopp, D., & Schedl, P. (2000). Sex determination in the drosophila germline is dictated by the sexual identity of the surrounding soma. Genetics, 155(4), 1741-1756.

Wawersik, M., Milutinovich, A., Casper, A.L., Matunis, E., Williams, B., and Van Doren, M. (2005). Somatic control of germline sexual development is mediated by the JAK/STAT pathway. Nature 436, 563-567.

Weng, R., Chin, J.S., Yew, J.Y., Bushati, N., and Cohen, S.M. (2013). miR-124 controls male reproductive success in Drosophila. eLife 2, e00640.

White, K. P., Hurban, P., Watanabe, T., & Hogness, D. S. (1997). Coordination of drosophila metamorphosis by two ecdysone-induced nuclear receptors. Science (New York, N.Y.), 276(5309), 114-117.

Whitworth, C., Jimenez, E., & Van Doren, M. (2012). Development of sexual dimorphism in the drosophila testis. Spermatogenesis, 2(3), 129-136. doi:10.4161/spmg.21780

Williams, T., et al. (2008). The regulation and evolution of a genetic switch controlling sexually dimorphic traits in Drosophila. Cell. 134, 610-623.

Williams, T.M., Selegue, J.E., Werner, T., Gompel, N., Kopp, A., and Carroll, S.B. (2008). The regulation and evolution of a genetic switch controlling sexually dimorphic traits in Drosophila. Cell 134, 610-623.

Xie, T. (2013). Control of germline stem cell self-renewal and differentiation in the drosophila ovary: Concerted actions of niche signals and intrinsic factors. Wiley Interdisciplinary Reviews.Developmental Biology, 2(2), 261-273. doi:10.1002/wdev.60

Yang, Y., Zhang, W., Bayrer, J. R., & Weiss, M. A. (2008). Doublesex and the regulation of sexual dimorphism in drosophila melanogaster: Structure, function, and mutagenesis of a female-

specific domain. The Journal of Biological Chemistry, 283(11), 7280-7292. doi:10.1074/jbc.M708742200

Yang, Y., Zhang, W., Bayrer, J.R., and Weiss, M.A. (2008). Doublesex and the regulation of sexual dimorphism in Drosophila melanogaster: structure, function, and mutagenesis of a female-specific domain. The Journal of biological chemistry 283, 7280-7292.

Yao, T. P., Forman, B. M., Jiang, Z., Cherbas, L., Chen, J. D., McKeown, M., . . . Evans, R. M. (1993). Functional ecdysone receptor is the product of EcR and ultraspiracle genes. Nature, 366(6454), 476-479. doi:10.1038/366476a0

Yi, W., and Zarkower, D. (1999). Similarity of DNA binding and transcriptional regulation by Caenorhabditis elegans MAB-3 and Drosophila melanogaster DSX suggests conservation of sex determining mechanisms. Development (Cambridge, England) 126, 873-881.

Yoder, J.H. (2012). Abdominal segment reduction: development and evolution of a deeply fixed trait. Fly 6, 240-245.

Yu, R. T., McKeown, M., Evans, R. M., & Umesono, K. (1994). Relationship between drosophila gap gene tailless and a vertebrate nuclear receptor tlx. Nature, 370(6488), 375-379. doi:10.1038/370375a0

Zarkower, D. (2013). DMRT genes in vertebrate gametogenesis. Current topics in developmental biology 102, 327-356.

Zeisig, D.T., Bittner, C.B., Zeisig, B.B., Garcia-Cuellar, M.P., Hess, J.L., and Slany, R.K. (2005). The eleven-nineteen-leukemia protein ENL connects nuclear MLL fusion partners with chromatin. Oncogene 24, 5525-5532.

Zelhof, A. C., Yao, T. P., Chen, J. D., Evans, R. M., & McKeown, M. (1995). Seven-up inhibits ultraspiracle-based signaling pathways in vitro and in vivo. Molecular and Cellular Biology, 15(12), 6736-6745.

Zhang, W., Li, B., Singh, R., Narendra, U., Zhu, L., and Weiss, M.A. (2006). Regulation of sexual dimorphism: mutational and chemogenetic analysis of the doublesex DM domain. Molecular and cellular biology 26, 535-547.

Zhang, Y., Sturgill, D., Parisi, M., Kumar, S., and Oliver, B. (2007). Constraint and turnover in sex-biased gene expression in the genus Drosophila. Nature 450, 233-237.

Erin Alisa Jimenez

Johns Hopkins University ♦ 3400 North Charles Street ♦ Mudd Hall 302 ♦ Baltimore, MD 21218 ♦ Lab: (410) 516-4830 ♦ E-mail: ejimene4@jhu.edu

---

Education

Ph.D. in Cell, Molecular, and Developmental Biology
April 2015
Johns Hopkins University
Baltimore, Maryland

Bachelor of Science in Molecular, Cell, and Developmental Biology
June 2009
University of California, Los Angeles
Los Angeles, CA

Associate of Arts in General Sciences
June 2007
Santa Monica College
Santa Monica, CA

Graduate Research

NIH Ruth L. Kirschstein National Research Service Award Pre-doctoral Fellow
Johns Hopkins University, Department of Biology
2009-2015
Principle Investigator: Mark Van Doren
Thesis: “The role of doublesex in sexually dimorphic development in Drosophila”

- Identified genomic loci bound by the transcription factor Doublesex
- Discovered transcripts that change in response to Doublesex
- Determined roles of identified Doublesex targets in regulating sexual dimorphism
- Investigated sex-specific steroid hormone signaling in the gonad

Undergraduate Research Experiences

MARC U-STAR/Bridge Program
2007-2009
Department of Chemistry and Biochemistry
University of California, Los Angeles
Principle Investigator: Carla Koehler

- Investigated the significance of mitochondrial localization of AKT
- Investigated the role of miro in trafficking mitochondria in zebrafish
- Localized the mitochondrial carrier protein MTCH2 to the outer membrane in yeast

Undergraduate Research Program
June 2008-August 2008
Watson School of Biological Sciences at Cold Spring Harbor
Principle Investigator: David Jackson

- Investigated the regulatory process involved in plant morphogenesis using maize *abphyl1* and *Abphyl2* mutants

Publications

Research Papers:

Jimenez, E., and Van Doren, M. Doublesex regulates sex-specific niche formation via Ecdysteroid hormone signaling in the Drosophila gonad. (*In Preparation*).

Clough, E., **Jimenez, E**.*, Kim Y., Whitworth, C., Goodwin, M., Hempel, L., Pavlou, H., Chen, Z., Sturgill, D., Dale, R., Smith., H. E., Przytycka, T. M., Goodwin, S., Van Doren, M., and Oliver, B. Sex- and tissue-specific functions of Drosophila Doublesex transcription factor target genes. Developmental Cell, 31(6): 761-773, 2014 Dec. **Co-first author.*

Zaltsman, Y., Shachnai, L., Yivgi-Ohana, N., Schwarz, M., Maryanovich, M., Houtkooper, R. H., Maxime Vaz, F., De Leonardis, F., Fiermonte, G., Palmieri, F., Gillissen, B., Daniel, P, **Jimenez, E.**, Walsh, S., Koehler, C. M., Roy, S. S., Walter, L, Hajnóczky, G, Gross, A. MTCH2/MIMP is a major facilitator of tBID recruitment to mitochondria. Nature Cell Biology, 12(6): 553-62, 2010 May.

Reviews:

Whitworth, C., **Jimenez, E**., and Van Doren M. Development of sexual dimorphism in the Drosophila testis. Spermatogenesis 2:129-136; http://dx.doi.org/10.4161/spmg.21780, 2012 July.

Acknowledgements:

Lee B, Johnston R, Yang Y, Gallavotti A, Kojima M, Travencolo B, Costa L, Sakakibara H, Jackson D. Studies of aberrant phyllotaxy1 mutants of maize indicate complex interactions between auxin and cytokinin signaling in the shoot apical meristem. Plant Physiology, 150(1): 205-16, 2009 March.

Abstracts:

Talks:

Jimenez, E.*, and DSX Consortium. (2014) “A genomic analysis of doublesex targets and function in sex determination” presented at Rollins College (**Invited speaker*).

Posters:

Jimenez, E., and Van Doren, M. (2014) “Regulation of niche initiation by sex and hormonal signaling” presented at the Mid-Atlantic Society for Developmental Biology Meeting.

Jimenez, E., and DSX Consortium. (2014) “A genomic analysis of doublesex targets and function in sex determination” presented at the 55th Drosophila Research Conference.

Jimenez, E., and Van Doren, M. (2013) “The role of sex and hormonal regulation in niche initiation” presented at the Johns Hopkins University Symposium.

Jimenez, E., and Van Doren, M. (2012) “A genomics analysis of sex determination” presented at the Johns Hopkins University CMDB Recruitment Symposium.

Jimenez, E., and Van Doren, M. (2011) “The role of Drosophila doublesex in the development of a sexually dimorphic gonad” presented at the 52nd Drosophila Research Conference.

Jimenez, E., Walsh, S., Koehler, C. “Active Akt Localizes to the Intermembrane Space of Mammalian Brain Mitochondria” presented at the American Society for Cell Biology (ASCB) Annual Meeting (2008); UCLA May Science Poster Day (2008 and 2009); CAMP Statewide Symposium, Beckman Center of the National Academics of Sciences and Engineering, (2008).

Jimenez, E., Walsh, S., Gross, A., Koehler, C. (2008) “MTCH2/MIMP localizes to the mitochondrial outer membrane” presented at the Society for the Advancement of Chicanos and Native Americans in Science (SACNAS) National Conference.

Jimenez, E., Johnston, R., Jackson, D. (2008) “Determining the mechanisms controlling phyllotaxy in maize with *abphyl1* and *Abphyl2* mutants” presented at UCLA Minority Access to Research Careers (MARC) Undergraduate Research Symposium and at the Cold Spring Harbor Laboratory Undergraduate Research Symposium.

Teaching Experience

Mentor of Graduate Research Project	9/2013-12/2013
Laboratory Teaching Assistant, Developmental Biology	1/2011-6/2011
Teaching Assistant, Biochemistry Lecture	5/2012-6/2012
Head Teaching Assistant, Genetics Lecture	9/2011-12/2011
Laboratory Teaching Assistant, Genetics	9/2010-1/2011
Teaching Assistant, Genetics Lecture	9/2010-1/2011

Academic and Professional Honors

2013-2014 Recipient of the Ruth L. Kirschstein National Research Service Award for Individual Pre-doctoral Fellowships to Promote Diversity in Health Related Research (F31)

2011	DuPont Teaching Award for Developmental Biology, Johns Hopkins University
2009	Recipient of the Women for Change Award, UCLA
2009	The Elma Gonzalez Prize for superb achievement in the Minority Access to Research Careers Program, NIGMS
2008	Undergraduate Research Fellowship, Watson School of Biological Sciences at Cold Spring Harbor Laboratory
2008	Special Merit for Biological/Life Sciences Poster, CAMP Statewide Symposium, Beckman Center of the National Academics of Sciences and Engineering, UC Irvine
2008	UCLA Dean's Prize for Outstanding Research, UCLA Science Poster Day
2007	Undergraduate Research Fellowship, University of California, Los Angeles, Bridge Summer Research Program for Community College Students
2007-2009	Minority Access to Research Careers (MARC) Research Fellowship, UCLA
2005	Excellence in Freshman Biology, Mount St. Mary's College
2004-2005	Dean's List Gold, Mount St. Mary's College